



**UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA DE ALIMENTOS**

YASMIN LIMA BRASIL

**ESPECTROSCOPIA DE INFRAVERMELHO PRÓXIMO (NIRS) PARA
DETERMINAÇÃO DA QUALIDADE E FRESCOR DE OVOS DE CODORNA**

**NEAR INFRARED SPECTROSCOPY (NIRS) FOR DETERMINATION OF
QUALITY AND FRESHNESS OF QUAIL EGGS**

**CAMPINAS – SP
2021**

YASMIN LIMA BRASIL

**ESPECTROSCOPIA DE INFRAVERMELHO PRÓXIMO (NIRS) PARA
DETERMINAÇÃO DA QUALIDADE E FRESCOR DE OVOS DE CODORNA**

**NEAR INFRARED SPECTROSCOPY (NIRS) FOR DETERMINATION OF QUALITY
AND FRESHNESS OF QUAIL EGGS**

Dissertação apresentada à Faculdade de
Engenharia de Alimentos da Universidade
Estadual de Campinas como parte dos
requisitos exigidos para a obtenção do título de
Mestra em Engenharia de Alimentos.
Dissertation presented to the School of Food
Engineering of the University of Campinas in
partial fulfillment of the requirements for the
degree of Master, in Food Engineering.

Supervisor/Orientador: Prof. Dr. Douglas Fernandes Barbin

Este trabalho corresponde a versão final
da dissertação defendida pela aluna
Yasmin Lima Brasil, e orientada pelo
Prof. Dr. Douglas Fernandes Barbin.

CAMPINAS – SP

2021

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Faculdade de Engenharia de Alimentos
Claudia Aparecida Romano - CRB 8/5816

B736e Brasil, Yasmin Lima, 1996-
Espectroscopia de infravermelho próximo (NIRS) para determinação da qualidade e frescor de ovos de codorna / Yasmin Lima Brasil. – Campinas, SP : [s.n.], 2021.

Orientador: Douglas Fernandes Barbin.
Dissertação (mestrado) – Universidade Estadual de Campinas, Faculdade de Engenharia de Alimentos.

1. Espectrômetro portátil. 2. Ovos. 3. Armazenamento. 4. Aprendizado de máquina. I. Barbin, Douglas Fernandes. II. Universidade Estadual de Campinas. Faculdade de Engenharia de Alimentos. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Near infrared spectroscopy (NIRS) for determination of quality and freshness of quail eggs

Palavras-chave em inglês:

Portable spectrometer

Eggs

Storage

Machine learning

Área de concentração: Engenharia de Alimentos

Titulação: Mestra em Engenharia de Alimentos

Banca examinadora:

Douglas Fernandes Barbin [Orientador]

Ana Paula Ayub da Costa Barbon

Juliana Azevedo Lima Pallone

Data de defesa: 12-08-2021

Programa de Pós-Graduação: Engenharia de Alimentos

Identificação e informações acadêmicas do(a) aluno(a)

- ORCID do autor: <https://orcid.org/0000-0001-8698-2428>

- Currículo Lattes do autor: <http://lattes.cnpq.br/8196805399091170>

COMISSÃO EXAMINADORA

Prof. Dr. Douglas Fernandes Barbin

Orientador - Faculdade de Engenharia de Alimentos, UNICAMP

Prof^a. Dra. Ana Paula Ayub da Costa Barbon

Membro Titular – Centro Universitário Filadélfia, UNIFIL

Prof^a. Dra. Juliana Azevedo Lima Pallone

Membro Titular - Faculdade de Engenharia de Alimentos, UNICAMP

A Ata da defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa da Unidade.

DEDICATÓRIA

*Aos amores da minha vida Janaina Brasil, José
Bandeira, Isabelly Brasil, Ysadora Bandeira, Dariana
Bandeira, Josefa Bandeira, Djario Bandeira e
Flaviana Angelim.*

AGRADECIMENTOS

Primeiramente à Deus, por ter me concedido saúde e força para continuar a minha trajetória acadêmica e por estar constantemente presente em minha vida. Sem Ele, nada disso seria possível.

Aos meus pais Janaina Brasil e José Bandeira, por todo apoio, incentivo, carinho, força e amor em todos os momentos até aqui, principalmente os mais difíceis. As minhas irmãs Isabelly Brasil, Ysadora Bandeira e Dariana Bandeira, pelo companheirismo, amor e carinho. A minha vó Josefa Bandeira e aos meus tios Djario Bandeira e Flaviana Angelim, por sempre me ajudarem e acreditarem em mim. Amo vocês demais.

Ao meu orientador Douglas Fernandes Barbin e aos meus amigos Luis Jam Pier Cruz Tirado e Maria Lucimar da Silva Medeiros por todos os ensinamentos, oportunidades de crescimento, paciência, carinho, cuidado e apoio em todos os momentos da minha trajetória acadêmica e pessoal, por me estimularem a ser cada dia uma profissional melhor. Vocês são exemplos pra mim, quero levar a amizade de vocês pro resto da vida.

Aos meus amigos Karina Rodrigues, Raquel Sá, Renann Lopes, Heryca Nobrega e Anderson Formiga que contribuíram para que esse sonho se tornasse realidade, tenho um carinho imenso por cada um de vocês, agradeço por todo incentivo e todos os momentos compartilhados.

A toda a equipe do laboratório de Inovação em Alimentos (LINA), em especial aos colegas do grupo de pesquisa em Tecnologias Analíticas Emergentes pelas experiências vivenciadas e pelas contribuições para a execução deste trabalho.

A Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (projeto número 2015/24351-2) pela bolsa (2019/11896-1) subsidiada, a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Código Financeiro 001 e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo financiamento da pesquisa.

RESUMO

Os ovos são mundialmente consumidos por serem uma fonte acessível de nutrientes essenciais na alimentação humana. Após a ovoposição o ovo está sujeito a alterações físicas e químicas que podem resultar em perdas na qualidade interna. Os ovos de codorna possuem um valor nutricional superior aos ovos de galinha. A qualidade dos ovos é avaliada através de métodos físicos e químicos complexos, demorados, e invasivos, que acarretam grande perda econômica para a indústria. Neste contexto, técnicas não destrutivas como a espectroscopia no infravermelho próximo (NIRS) associada a ferramentas quimiométricas são uma excelente alternativa para a análise e monitoramento de parâmetros em alimentos, pois permite realizar medições simples e rápidas, determinação simultânea de multicomponentes, apresentando baixo custo. Este projeto teve como objetivo realizar um estudo sobre o potencial de um espectrômetro portátil para analisar a qualidade e o frescor de ovos de codorna durante o armazenamento através da determinação da Unidade Haugh (HU), Índice de Gema (YI) e Índice de Qualidade do Ovo (EQI). Pré-tratamentos matemáticos como centralização na média (mean-center), cálculo da derivada (SG) e variável normal padrão (SNV) foram aplicados para eliminar efeitos externos aos espectros. A análise de componentes principais (PCA) foi empregada como uma ferramenta exploratória dos espectros. Os modelos preditivos elaborados através da regressão por máquina de vetor de suporte (SVMR) apresentaram desempenho superior a regressão por mínimos quadrados parciais (PLSR) para predição de HU, IG e EQI em ovos de codorna intactos, demonstrando boa capacidade de predição, com RPD de 2,0 - 2,5 e RER >10. A análise discriminante de mínimos quadrados parciais (PLSDA) e a classificação por máquina de vetor de suporte (SVMC) foram utilizadas para discriminação entre ovos de codorna frescos e não-frescos, demonstrando serem eficientes para classificar corretamente mais de 80% das amostras, com acurácias entre 79,3 – 85,4% e 79,6 – 85,3%, respectivamente. Através dos resultados obtidos neste estudo foi possível comprovar a efetividade da espectroscopia NIR, em combinação com a análise multivariada, para avaliação da qualidade e do frescor de ovos de codorna mediante abordagens quantitativas e qualitativas.

Palavras-chave: Espectrômetro portátil, ovos, frescor, armazenamento, aprendizado de máquina.

ABSTRACT

Eggs are consumed worldwide as an accessible source of essential nutrients in human nutrition. After laying, the egg is subject to physical and chemical changes that can result in losses in internal quality. Quail eggs have a higher nutritional value than chicken eggs. Egg quality is evaluated through complex, time-consuming, and invasive physical and chemical methods, which cause great economic loss for the industry. In this context, non-destructive techniques such as near infrared spectroscopy (NIRS) associated with chemometrics are an excellent alternative for the analysis and monitoring of parameters in foods, as this allows simple and fast measurements, simultaneous determination of multicomponents, with low cost. This project aimed to carry out a study on the potential of a portable spectrometer to analyze the quality and freshness of quail eggs during storage by determining the Haugh Unit (HU), Yolk Index (YI) and Index of Egg Quality (EQI). Mathematical pre-treatments such as mean centering, Savitzky-Golay derivative calculation (SG) and standard normal variable (SNV) were applied to the data to facilitate the interpretation of the spectra. Principal component analysis (PCA) was used as an exploratory tool for spectral data. The predictive models elaborated by means of support vector machine regression (SVMR) performed better than those elaborated by partial least squares regression (PLSR) for predicting HU, YI and EQI in intact quail eggs, demonstrating good predictive capacity, with RPD of 2.0 - 2.5 and RER >10. Partial least squares discriminant analysis (PLSDA) and support vector machine classification (SVMC) were used to discriminate between fresh and non-fresh quail eggs, with classification accuracy over than 80% of the samples, with accuracy between 79.3 – 85.4% and 79.6 – 85.3%, respectively. Results proved the effectiveness of NIR spectroscopy, in combination with multivariate analysis, to assess the quality and freshness of quail eggs through quantitative and qualitative approaches.

Keywords: Portable spectrometer, eggs, freshness, storage, machine learning.

SUMÁRIO

CAPÍTULO 1 – INTRODUÇÃO GERAL	10
1.1 INTRODUÇÃO GERAL	11
CAPÍTULO 2 – REVISÃO DE LITERATURA	13
2.1. REVISÃO DE LITERATURA	14
2.1.1. OVO NA DIETA HUMANA: OVOS DE CODORNA	14
2.1.2. DETERMINAÇÃO DO FRESCOR DE OVOS	15
2.1.3. MÉTODOS NÃO DESTRUTIVOS	17
2.1.4. ANÁLISE MULTIVARIADA	18
2.1.4.1. PRÉ-TRATAMENTO	19
2.1.4.2. ANÁLISE EXPLORATÓRIA	20
2.1.4.3. ANÁLISE DE REGRESSÃO	21
2.1.4.4. ANÁLISE DE CLASSIFICAÇÃO	22
CAPÍTULO 3 – FAST AND ONLINE FRESHNESS ESTIMATION OF QUAIL EGGS USING PORTABLE NIR ESPECTROMETER AND MACHINE LEARNING	24
3.1. INTRODUCTION	27
3.2. MATERIAL AND METHODS	29
3.2.1. EGG SAMPLE	29
3.2.2. SPECTRA ACQUISITION	30
3.2.3. EGG QUALITY INDEXES	31
3.2.3.1. HAUGH UNIT (HU)	31
3.2.3.2. YOLK INDEX (YI)	31
3.2.3.3. EGG QUALITY INDEX (EQI)	31
3.2.4. DATA ANALYSIS	32
3.2.4.1. EXPLORATORY ANALYSIS	32
3.2.4.2. PREDICTION OF EGGS QUALITY PARAMETERS	33
3.2.4.3. EGG FRESHNESS CLASSIFICATION	34
3.2.5. STATISTICAL ANALYSIS	35
3.3 RESULTS AND DISCUSSION	35

3.3.1. FRESHNESS OF EGGS DURING STORAGE	35
3.3.2. SPECTRAL AND EXPLORATORY ANALYSIS	37
3.3.3. PREDICTION OF EGGS QUALITY INDEXES	40
3.3.4. EGG FRESHNESS GRADING	46
3.4. CONCLUSION	52
ACKNOWLEDGMENT	52
CONFLIC OF INTEREST	52
REFERENCES	53
CAPÍTULO 4 – CONCLUSÃO GERAL	59
CAPÍTULO 5 – REFERÊNCIAS	61

CAPÍTULO 1:
INTRODUÇÃO GERAL

1.1. INTRODUÇÃO GERAL

Os ovos são mundialmente consumidos por serem uma fonte acessível de nutrientes essenciais, como antioxidantes, aminoácidos, proteínas de alta qualidade e ácidos graxos de fácil digestão, além de representarem ingredientes importantes para a indústria alimentícia devido às suas propriedades tecnológicas como, formação de espuma, gelificação e emulsificação (LOFFREDI; GRASSI; ALAMPRESE, 2021). Diferentes espécies de aves possuem características únicas de ovo, levando a valores nutricionais variados. Apesar dos ovos de codorna representarem quase um quinto do tamanho dos ovos de galinha, estes possuem um valor nutricional superior (BAO et al., 2020; SUN et al., 2019).

No entanto, após a ovoposição o ovo está sujeito a alterações físicas e químicas no albúmen e na gema que resultam em perdas na qualidade interna. Essas alterações estão relacionadas a idade, condições ambientais, processo de armazenamento, venda e processamento dos ovos (PIRES et al., 2019). O grau de frescor é um dos índices mais importantes para avaliar a qualidade dos ovos, no entanto torna-se difícil esta identificação apenas pela aparência, havendo a necessidade da utilização de métodos físicos e químicos complexos, demorados e invasivos, que acarretam grande perda econômica para a indústria (REN et al., 2020).

Neste contexto, técnicas não destrutivas como a espectroscopia no infravermelho próximo (NIRS) associada a ferramentas quimiométricas tem se mostrado uma excelente alternativa para a análise e monitoramento de parâmetros em alimentos (ABU-KHALAF; HMIDAT, 2020). Quando empregada esta técnica permite realizar medições simples e rápidas, detectando com segurança a composição interna e discriminando com precisão as mudanças de qualidade durante o armazenamento, além de apresentar baixo custo e permitir a determinação simultânea de multicomponentes, podendo ser totalmente automatizada (ROSSI; LOZANO, 2020; TEYE et al., 2020).

A espectroscopia NIR consiste em um método de espectroscopia vibracional que utiliza radiação eletromagnética entre 750 e 2500 nm, localizado entre o infravermelho e a região visível, variando com a ordem do sobretom, natureza e força da ligação. As variações nesse intervalo do espectro estão relacionadas principalmente a ligações moleculares como C-H, O-H, S-H e N-H, que são os grupos de átomos estruturais da maioria dos compostos orgânicos (MARQUES et al., 2017). Os desafios para utilização dessa técnica incluem matrizes heterogêneas, uma vez que esta fornece informações pontuais de uma área relativamente pequena da amostra, e desenvolvimento dos modelos quimiométricos necessários para

contornar o problema da falta de seletividade da informação analítica contida nos espectros, facilitando a interpretação dos resultados (RIGOLLE et al., 2015).

A absorção de radiação na região do infravermelho próximo (NIR) por moléculas orgânicas dá origem a espectros bastante complexos, neste contexto uma ampla gama de algoritmos de regressão e classificação foram desenvolvidos para analisar informações do espectro, selecionando os comprimentos de onda mais informativos (MARQUES et al., 2017). A análise de componentes principais (PCA) é frequentemente empregada como uma ferramenta exploratória dos dados espectrais, transformando um conjunto de dados altamente correlacionados em um grupo menor de componentes não correlacionados, reduzindo a totalidade de dados original (ESQUERRE et al., 2020). Dentre os diversos métodos matemáticos comumente utilizados para prever parâmetros quantitativos e qualitativos de amostras, estão a regressão por mínimos quadrados parciais (PLS) e máquina de vetor de suporte (SVM) (PULLANAGARI; LI, 2020).

Neste contexto, a espectroscopia NIR aliada á métodos de análise multivariada, representam um grande potencial para o monitoramento do frescor e de parâmetros internos de qualidade de ovos de codorna durante a estocagem. Diante disso, o objetivo deste estudo foi investigar o potencial de um espectrômetro portátil, em combinação com aprendizado de máquina, para determinação de parâmetros associados a qualidade de ovos de codorna, a fim de classificá-los de acordo com o seu frescor, destacando as melhores características para o seu processamento.

CAPÍTULO 2:

REVISÃO DE LITERATURA

2.1. REVISÃO DE LITERATURA

2.1.1. OVO NA DIETA HUMANA: OVOS DE CODORNA

Nos últimos anos, especialmente durante a pandemia do Covid-19, a produção e o consumo de ovos aumentaram em todo o mundo devido a diversidade de seus nutrientes junto com sua alta digestibilidade e seu preço acessível, representando uma fonte proteica de alta qualidade. Além das quantidades de proteína os ovos são uma fonte alimentar rica em aminoácidos essenciais (lisina), ácidos graxos, diferentes vitaminas, incluindo A, B, D, E e K e minerais como ferro, cálcio, magnésio, selênio, sódio, zinco e fósforo (SHARAF EDDIN; IBRAHIM; TAHERGORABI, 2019).

Os ovos constituem uma das principais fontes de nutrientes na dieta humana, sendo considerados um dos alimentos mais completos, contendo inúmeros compostos bioativos, que podem ser de grande interesse na prevenção de muitas doenças, o que o torna um alimento funcional (RÉHAULT-GODBERT; GUYOT; NYS, 2019). Ovos de diferentes espécies avícolas formaram suas propriedades únicas, desde o tamanho do ovo até a composição nutricional e funcional, devido a processos de evolução divergentes (SUN et al., 2019). Os ovos de codorna possuem uma estrutura morfológica semelhante aos ovos de galinha, diferindo apenas na composição e proporção. Do ponto de vista nutricional, os ovos de codorna são mais nutritivos do que outros ovos porque são ricos em antioxidantes, minerais e vitaminas (BAO et al., 2020). Pesquisas também demonstraram que o teor de proteína da clara de ovo de codorna é superior ao da clara de ovo de galinha (MIGUEL et al., 2005). Na tabela abaixo temos uma comparação da composição dos ovos mais consumidos no Brasil e no mundo (Tabela 1).

Tabela 1. Composição química dos ovos de galinha, codorna, pata, perua e gansa (100 gramas).

Espécie	Proteína (g)	Lipídeos (g)	Carboidratos (g)	Umidade (g)	Calorias (cal)
Galinha	12,14	11,15	1,20	74,57	155
Codorna	13,05	11,09	0,41	74,35	160
Pata	12,81	13,77	1,45	70,83	185

Perua	13,68	11,88	1,15	72,50	168
Gansa	13,85	13,27	1,35	70,43	185

Fonte: (BARBOSA, 2013).

Mudanças nos hábitos alimentares da população favoreceram o aumento do consumo de ovos de codorna em todo o mundo, devido ao seu sabor único, excelente valor nutricional e propriedades funcionais, sendo considerada uma proteína animal de alto valor biológico, atendendo às necessidades humanas diárias principalmente de aminoácidos essenciais (BRAGAGNOLO; RODRIGUEZ-AMAYA, 2003; HE et al., 2021; NASR et al., 2019). O desenvolvimento acentuado da coturnicultura também está relacionado ao fato das codornas serem aves resistentes a doenças, apresentarem alta produtividade (média de 300 ovos/ano), demandarem pouco espaço e utensílios para sua manipulação, exigindo baixo custo de investimento e criação e, conseqüentemente, um rápido retorno financeiro (BRAGAGNOLO; RODRIGUEZ-AMAYA, 2003; MÓRI et al., 2005). Isso resultou em um aumento significativo na produção de ovos de codorna. Em 2019, a produção de ovos de codorna no Brasil foi estimada em 3.786,95 bilhões de ovos, sendo concentrada principalmente na região sudeste (67,32%) (IBGE, 2020). Estudos recentes demonstram que o consumo regular de ovos de codorna pode combater uma variedade de doenças, aumentar a imunidade, promover a saúde da memória, aumentar a atividade cerebral, estabilizar o sistema nervoso, remover toxinas e metais pesados, melhorando o metabolismo (LIU et al., 2020; OLADIPO et al., 2020).

2.1.2. DETERMINAÇÃO DO FRESCOR DE OVOS

A qualidade geral e frescor dos ovos são cruciais tanto para avicultores quanto para consumidores. As medições da qualidade do ovo apresentam duas divisões principais: características externas e internas. Os atributos de qualidade externos, como peso do ovo e condição da casca, são significativos para a aceitação do consumidor, pois na maioria das vezes, considera-se que a qualidade interna tem uma correlação direta com suas características externas, enquanto os parâmetros internos são fundamentais para a indústria de processamento, desempenhando um papel muito significativo na elaboração de diversos produtos (EL-TARABANY, 2016).

Os ovos são altamente perecíveis e podem perder rapidamente sua qualidade interna devido à porosidade da casca que permite trocas gasosas com o meio externo, facilitando a perda de água e CO₂, sendo facilmente deteriorados durante a produção, processamento, venda e distribuição. A taxa na qual essas mudanças ocorrem no albúmen e na gema depende principalmente do intervalo entre o tempo de coleta e o tempo de consumo dos ovos, resultando em alterações no frescor, sabor, palatabilidade e, conseqüentemente, em perdas econômicas para a indústria. Assim, o frescor desempenha um papel muito significativo na percepção da qualidade pelo consumidor (KUMBÁR et al., 2015; SARI; TILKI; SAATCI, 2016).

De acordo com a legislação brasileira (BRASIL, 1990), a refrigeração dos ovos nos pontos de venda é opcional e, portanto, não ocorre em condições práticas. De acordo com (OLIVEIRA; OLIVEIRA, 2013), no Brasil, 92% dos ovos comercializados *in natura* são expostos à temperatura ambiente e resfriados apenas na casa do consumidor, fazendo com que os nutrientes contidos no interior dos ovos sejam transformados rapidamente em substâncias impróprias para o consumo humano, depreciando portanto sua qualidade interna (MARINHO, 2011). Existem vários indicadores para estimar o frescor e a qualidade interna dos ovos de galinha como, altura do albúmen, unidade Haugh, pH do albúmen e da gema, índice de gema, perda de peso e tamanho da camada de ar (KUMBÁR et al., 2015). No entanto, não existe legislação específica para regulamentação do frescor de ovos de codorna, o que acaba dificultando sua aplicação tecnológica e funcional (SUN et al., 2019).

A Unidade Haugh (HU) é o índice internacional utilizado para avaliar o frescor e a qualidade interna dos ovos da maioria das espécies avícolas. HU descreve as alterações sofridas pelo albumén, principalmente devido a mudanças no complexo ovomucina-lisozima. O Índice de Gema (YI) também é bastante utilizado, pois indica o achatamento da gema decorrente da transferência de água do albúmen para a gema através da membrana vitelínea. Quanto maior forem ambos os parâmetros, maior a qualidade interna e o frescor dos ovos (MALFATTI et al., 2021). O pH aumenta rapidamente após a postura, favorecendo a liquefação do albumén devido o aumento da acidez no meio. Já a perda de peso e o crescimento da camada de ar são causados principalmente pela difusão de água através da casca e pela separação da membrana da casca, aumentando significativamente com o aumento do tempo e temperatura de armazenamento dos ovos (LOFFREDI; GRASSI; ALAMPRESE, 2021).

Como já discutido anteriormente essas reações físico-químicas estão intimamente relacionadas às trocas gasosas com o meio externo, resultando em perdas na qualidade interna. Apesar desses parâmetros descreverem de maneira eficiente o frescor e a qualidade interna dos ovos de galinha, a coturnicultura necessita de um parâmetro mais abrangente que detalhe a influência do período de armazenamento tanto nas características do albúmen como da gema. O Índice de Qualidade do Ovo (EQI), pode ser uma ferramenta valiosa para atingir esse objetivo, pois relaciona características externas, como o peso do ovo, e características internas, como a altura do albúmen e da gema, sendo proporcional a HU (NARUSHIN; ROMANOV; GRIFFIN, 2021). Na figura 1 estão representadas as regiões de medição dos parâmetros físicos dos ovos de codorna.

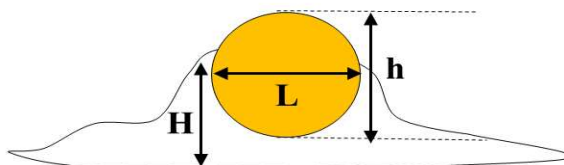


Figura 1. Região de medição dos parâmetros físicos. H: altura do albúmen espesso (mm), h: altura da gema (mm), L: largura da gema (mm). (Fonte: autor)

2.1.3. MÉTODOS NÃO DESTRUTIVOS

Diferentes métodos automatizados e manuais são empregados para avaliar o frescor e a qualidade dos ovos, uma vez que esta avaliação é essencial para classificá-los em relação à sua segurança e valor nutricional (SHARAF EDDIN; IBRAHIM; TAHERGORABI, 2019). No entanto a maioria desses métodos só podem ser aplicados a nível de laboratório, desconsiderando a urgência por programas de estratégias analíticas eficientes devido à crescente demanda da indústria alimentícia. A utilização de técnicas não destrutivas para determinação simples, econômica, rápida e segura da qualidade e frescor de produtos alimentícios tem aumentado nos últimos anos. Entre as tecnologias emergentes na região do infravermelho, temos a espectroscopia no infravermelho próximo (NIR), médio (MIR), distante (FIR) e visível (VISNIR), onde a maior parte das aplicações em análises quantitativas e qualitativas tem sido realizadas na região do infravermelho próximo e médio. Essas técnicas têm sido frequentemente utilizadas como

ferramentas analíticas na determinação de constituintes em alimentos, permitindo ampla aplicação, pouco ou nenhuma preparação de amostra e redução do tempo e dos custos de análise, com tecnologias ecologicamente corretas (HE et al., 2021; LOFFREDI; GRASSI; ALAMPRESE, 2021; MORGANO et al., 2007).

A espectroscopia no infravermelho próximo (NIRS), juntamente com algoritmos quimiométricos, tem sido amplamente estudada como alternativa potencial para abordagens químicas e físicas nas indústrias alimentícia, agrícola, farmacêutica e combustíveis. A espectroscopia NIR é um tipo de espectroscopia vibracional, no espectro eletromagnético entre 780 e 2500 nm ($12.800 - 4.000 \text{ cm}^{-1}$). Esta técnica espectroscópica foi desenvolvida com base na interação entre a luz e os componentes químicos dos materiais, baseando-se no fato de que as ligações químicas das substâncias possuem frequências de vibração específicas, as quais correspondem à vibração das ligações moleculares. O espectro obtido é composto por bandas associadas aos sobretons detectáveis e combinações das vibrações moleculares envolvendo os grupos químicos OH, NH, CH e SH, que são os grupos de átomos estruturais da maioria dos compostos orgânicos (MORGANO et al., 2007; SOHAIB ALI SHAH et al., 2020; TEYE et al., 2020).

Avanços tecnológicos auxiliaram na miniaturização dos espectrômetros, permitindo a criação de dispositivos portáteis e manuais, de baixo custo e fácil utilização, tornando a espectroscopia NIR facilmente aplicável em diversas linhas de processo. O desempenho dessas unidades na avaliação de parâmetros de qualidade foi relatado na literatura em diversos produtos alimentícios, incluindo óleos, gordura trans, estévia, mel, suco vegetal, leite em pó, entre outros. Diferentes estudos baseados na espectroscopia NIR demonstraram a efetividade desta técnica para avaliação da qualidade e do frescor de ovos de galinha (GIUNCHI et al., 2008; LIU et al., 2020; PUERTAS; VÁZQUEZ, 2020), provando ser um método de custo acessível e não destrutivo capaz de identificar e quantificar constituintes físicos em bandas espectrais específicas para uma resposta rápida sobre a qualidade e frescor do ovo.

2.1.4. ANÁLISE MULTIVARIADA

A espectroscopia NIR fornece um conjunto de dados com uma grande quantidade de variáveis espectrais altamente correlacionadas que precisam de um procedimento eficaz para melhorar a falta de seletividade e extrair informações analíticas úteis e

significativas dos espectros adquiridos, fazendo-se necessário a utilização de métodos de análise multivariada de dados, também chamada de quimiometria (MENDOZA et al., 2020).

Ferramentas quimiométricas que envolvem métodos estatísticos e matemáticos multivariados permitem organizar o conjunto de dados de uma matriz e modelar propriedades químicas e físicas dos dados espectrais adquiridos, possibilitando determinações quantitativas e qualitativas. Essas ferramentas também proporcionam a detecção de alterações entre os tratamentos e analitos, análise simultânea de inúmeras variáveis, melhora da exatidão de previsão dos parâmetros de interesse, facilitando também a detecção de padrões inerentes aos dados. Assim é possível discriminar grupos e identificar quais fatores causam essa discriminação, caracterizando melhor um determinado conjunto de dados. Portanto, a quimiometria permite que esses métodos multivariados complexos tornem-se práticos para utilização em análises de rotina na indústria (FERREIRA, 2020).

2.1.4.1. PRÉ – TRATAMENTO

Primeiramente, a necessidade de pré-tratamento espectral deve ser considerada, pois esta é uma etapa fundamental para facilitar a interpretação dos espectros e minimizar ruídos ou outras fontes de variabilidade não informativas. As técnicas de pré-tratamento mais utilizadas no campo espectroscópico são a centralização na média (mean-center), correção de dispersão multiplicativa (MSC), correção de linha de base (BL), variável normal padrão (SNV), normalização (NL) e derivadas espectrais (SG) (LOFFREDI; GRASSI; ALAMPRESE, 2021; MORGANO et al., 2007).

O pré-processamento dos dados através da centralização na média consiste em calcular a média das intensidades para cada comprimento de onda e subtrair cada uma das intensidades do valor médio (BALCEROWSKA-CZERNIAK et al., 2020). Transformações matemáticas como, os métodos de correção de dispersão MSC, SNV, BL e NL, visam reduzir a variabilidade física entre os dados espectrais devido ao fenômeno de espalhamento da luz e interferências causadas pelo tamanho das partículas e diferenças de densidade das amostras. Já o cálculo da primeira e segunda derivada, com filtro Savitzky-Golay (SG), permite remover efeitos aditivos e multiplicativos, corrigir deslocamento e inclinação da linha de base e melhorar a separação de sinais não totalmente sobrepostos (DE SOUZA; POPPI, 2012). A seleção do método de pré-

tratamento mais adequado depende da origem dos dados espectrais e do objetivo final do estudo. Por este motivo, a aplicação de diferentes pré-tratamentos e combinações destes deve ser considerada para avaliar qual deles melhora o desempenho dos modelos (LOFFREDI; GRASSI; ALAMPRESE, 2021).

2.1.4.2 ANÁLISE EXPLORATÓRIA

A análise de componentes principais (PCA) é uma manipulação matemática não supervisionada aplicada aos dados das matrizes com a principal finalidade de representá-los de forma compacta, em função da existência de correlação entre as variáveis originais, sem perda de informações significativas. Ao reduzir o número de variáveis, a PCA permite observar de forma mais clara a tendência de distribuição das amostras, relações entre variáveis e classes, possíveis agrupamentos e a presença de amostras anômalas (outliers), que podem estar relacionadas a falhas no processo de aquisição dos espectros (SABIN; FERRÃO; FURTADO, 2004; VALDERRAMA et al., 2015).

Quando empregada esta técnica multivariada consiste em identificar os fatores subjacentes, denominadas componentes principais (CP), que são combinações lineares das variáveis originais do conjunto de espectros (DE SOUZA; POPPI, 2012). A obtenção desses CP é realizada a partir da decomposição da matriz (X), que corresponde ao conjunto original dos dados, em uma matriz de escores (T), que são as coordenadas das amostras em um novo sistema de eixos, e uma matriz de pesos (P), que contém as informações do peso de cada variável original na formação dos novos eixos, conforme descrito pela seguinte equação (MARETTO, 2011).

$$X = TP$$

Desta forma, a PCA decompõe a matriz de dados originais em uma soma de matrizes, em que a primeira componente principal (PC1) detém máxima variância do conjunto de dados, enquanto que a segunda componente principal (PC2) é traçada ortogonalmente à primeira, contendo informações que não foram explicadas pela PC1, de forma que cada componente principal é responsável pela fração sucessiva de variância dos dados e assim por diante. Sendo assim a maior fração da variância é descrita nos primeiros CP, permitindo a visualização dos dados através do gráfico de scores de um CP versos o outro. O número de componentes define o grau de complexidade do modelo,

cujo ajuste deve permitir a melhor previsão possível dos parâmetros de interesse das amostras (BALCEROWSKA-CZERNIAK et al., 2020; SABIN; FERRÃO; FURTADO, 2004).

2.1.4.3. ANÁLISES DE REGRESSÃO

Para determinações quantitativas é necessário a construção de regressões múltiplas, através de técnicas quimiométricas, que têm por finalidade desenvolver modelos empíricos ou semi-empíricos derivados dos dados, que permitam estabelecer uma relação matemática entre o conjunto de espectros gerados e os parâmetros químicos ou físicos obtidos a partir de um método de referência para quantificação de um composto específico (MORGANO et al., 2007).

Para construção de um modelo de regressão multivariado, a espectroscopia NIR necessita de um conjunto de amostras de calibração e validação. A calibração consiste numa etapa de treinamento onde o método de regressão é estabelecido através de uma série de operações que visam à obtenção de um modelo matemático que estime as características de interesse a partir das medidas instrumentais, relacionando as variáveis dependentes (Y) e independentes (X) (GODOY, 2010). Após a construção do modelo de calibração, este deve ser validado por meio de um conjunto de amostras que não foram incluídas na calibração, que são as amostras do conjunto de validação. Esta fase, também conhecida como teste, tem como principal objetivo determinar a real capacidade preditiva do modelo proposto para estimar as propriedades requeridas a partir dos espectros, validando o modelo (MARQUES et al., 2017).

A partir disso, é possível observar e analisar as figuras de mérito para comparação entre os valores previstos pelo modelo e os valores previamente obtidos através de uma metodologia padrão, assegurando a aplicabilidade do modelo (BEDIN et al., 2021). Vários algoritmos lineares e não lineares podem ser utilizados para construção dos modelos de regressão, por exemplo, Regressão por Mínimos Quadrados Parciais (PLSR) e Regressão de Máquina de Vetor de Suporte (SVMR), sendo o PLSR um dos métodos de regressão mais empregados na calibração multivariada, em função da sua facilidade de elaboração e interpretação (MORGANO et al., 2007; TEYE et al., 2020).

O PLS baseia-se em modelos lineares e é utilizado quando existe uma relação linear entre o sinal analítico (espectro) e a propriedade a ser determinada. Os fatores/compostos estimados no modelo PLS são chamados variáveis latentes (VLs), sendo que cada variável latente é obtida relacionando-se a matriz de dados das variáveis de entrada/preditoras (X) com a o vetor contendo os valores das variáveis de saída/respostas (Y), determinados pelo método

de referência. A escolha do número de variáveis latentes que devem ser incluídas no modelo e o erro médio são feitas através da validação cruzada, na qual uma amostra ou grupo do conjunto de calibração é retirada de cada vez e prevista pelo modelo construído com as demais (HE et al., 2021; NIAZI; GOODARZI; YAZDANIPOUR, 2008). No entanto, o PLS nem sempre é a melhor opção, especialmente em situações que carregam não linearidades intrínsecas, fazendo-se necessário um modelo não linear (MARQUES et al., 2017).

A máquina de vetor de suporte (SVM) é um método de aprendizado de máquina fundamentado na Teoria do Aprendizado Estatístico, que pode modelar tanto relações lineares quanto não lineares, que pode ser mais adequada para análises quantitativas. Este algoritmo apresenta uma gama de aplicações envolvendo problemas de regressão e classificação, apresentando desempenho superior aos algoritmos de referência na presença de não-linearidades (HE et al., 2021; TERAMOTO et al., 2020). O SVM consiste em elaborar classificadores lineares e não lineares, utilizando linhas de classificação e hiperplanos para separar os dados linearmente em classes. Os dados relevantes para classificação e os pontos que ajudam a construir o hiperplano são chamados de vetores de suporte (SVs), sendo definidos do mesmo modo para ambos os casos. Os SVs são as amostras que estão mais próximas dos hiperplanos e das linhas de classificação e são, portanto as mais difíceis de classificar (TERAMOTO et al., 2020).

2.1.4.4. ANÁLISES DE CLASSIFICAÇÃO

As abordagens de classificação mais comuns são divididas em métodos discriminantes e de modelagem de classe, conferindo cada amostra a uma classe previamente estabelecida, através das informações espectrais (KEMSLEY; DEFERNEZ; MARINI, 2019). Um dos métodos discriminantes mais utilizado em abordagens espectroscópicas é a Análise Discriminante de Quadrados Mínimos Parciais (PLSDA), em função da sua simplicidade e maturidade como técnica analítica. Nos últimos anos, algoritmos mais complexos como a Classificação de Máquina de Vetor de Suporte (SVMC), foram comumente aplicados em estudos envolvendo a espectroscopia NIR, devido à sua capacidade de generalização para conseguir a separação de diferentes classes, evitando o sobreajuste, além de permitir a modelagem não linear, ampliando a capacidade de discriminação dos modelos de classificação (LOFFREDI; GRASSI; ALAMPRESE, 2021).

O PLSDA e o SVMC são métodos discriminantes supervisionados que requerem a priori informações categóricas para fins de classificação. Nesses métodos as amostras do

conjunto de calibração são usadas para definir os limites que particionam o espaço multidimensional em várias regiões, cada uma atribuída a uma classe. Em seguida, cada amostra do conjunto de previsão é atribuída à classe que tem a probabilidade máxima de pertencer. A diferença entre os métodos discriminantes está na maneira de como estimam a probabilidade de pertencimento à classe (LOFFREDI; GRASSI; ALAMPRESE, 2021).

No PLSDA a discriminação é baseada na identificação de variáveis latentes (VLs) não correlacionadas que permitam a maior separação entre os grupos de amostras, no entanto estas VLs não fornecem 100% das informações analíticas (Santos et al., 2021). A otimização da quantidade de VLs é essencial para generalização do modelo PLSDA, pois quando definido menos VLs do que o valor ideal, obtém-se um problema de subajuste, já se o número for maior, surge a questão do sobreajuste (BEDIN et al., 2021). O PLSDA calcula um valor limite, entre os valores previstos, sendo que valores acima deste limite indicam que a amostra pertence à classe modelada e valores abaixo deste limite indicam que a amostra não pertence à classe modelada (GODOY, 2010).

Já o SVMC usa todas as variáveis para construir um hiperplano de separação e selecionar os vetores de suporte em um espaço multidimensional, através do mapeamento dessas variáveis pela função de Kernel (DEVOS et al., 2009). A função Kernel representa a relação entre os dados de entrada (X-matriz de dados) e a propriedade de saída a ser modelada (Y-vetor resposta) num espaço de alta dimensão, onde a solução pode ser apresentada como um problema linear. A função Kernel mais utilizada para modelagem de dados espectroscópicos é a função de base radial (RBF), que consiste em uma medida não linear de similaridade. Uma maior distância entre o hiperplano e os vetores de suporte garante um melhor desempenho do modelo (BARRAGÁN et al., 2021; MARETTO, 2011).

CAPÍTULO 3:

FAST AND ONLINE FRESHNESS ESTIMATION OF QUAIL EGGS USING PORTABLE NIR SPECTROMETER AND MACHINE LEARNING

The results of this chapter are published in *Food Control*

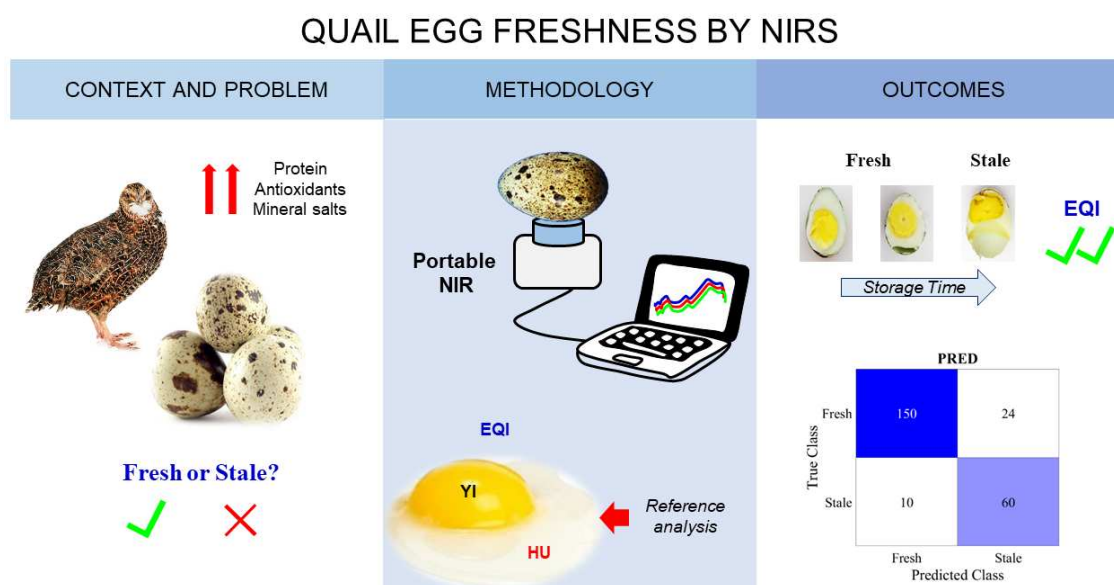
Fast and online freshness estimation of quail eggs using portable NIR spectrometer and machine learning

Yasmin Lima Brasil^a; J.P. Cruz-Tirado^a; Douglas Fernandes Barbin^{a,*}

Department of Food Engineering, University of Campinas, Rua Monteiro Lobato, 80, Cidade Universitária, Campinas, SP, 13083-862, Brazil

Corresponding author: dfbarbin@unicamp.br

Graphical abstract



Abstract

Quail eggs are one of the main natural sources of essential nutrients, presenting high amounts of protein, antioxidants, calcium, iron and phosphorus. However, its quality assessment demands laborious methods and chemicals, and there is currently no standard method to quantify its freshness. This work aimed to investigate the performance of a portable NIR spectrometer, in combination with machine learning, to estimate the freshness of quail eggs. Since there is no standard index to classify quail eggs, we compared Haugh Unit (HU), Yolk Index (YI) and the Egg Quality Index (EQI) as reference methods. Partial Least Squares Regression (PLSR) and Support Vector Machine Regression (SVMR) were used to build prediction models, and Partial Least Squares-Discriminant Analysis (PLSDA) and Support Vector Machine Classification (SVMC) for the development of classification models. For the first time, we demonstrated that EQI, which is a parameter that measures egg freshness according to the quality of the yolk and the albumen, is the best way to express the freshness

of quail eggs. The best prediction models were obtained for YI and EQI, using SVMR, with $RPD = 2.0 - 2.5$ and $RER > 10$, indicating good predictive capacity. PLSDA and SVMC models showed similar performance, correctly classifying more than 80% of the samples. The results obtained demonstrate the potential of portable NIR spectrometer for monitoring quail eggs freshness during storage.

Keywords: Near infrared spectroscopy, chemometrics, shelf life, data mining.

List of Abbreviations and Symbols

HU – Haugh unit

YI – Yolk index

EQI – Egg quality index

PLSR – Partial least squares regression

PLSDA – Partial least square discriminant analysis

RMSEC – Root mean square error of calibration

RMSECV – Root mean square error of cross validation

RMSEP – Root mean square error of prediction

RPD – ratio of standard deviation

RER – error reach rate

R^2_{CAL} – coefficient of determination of calibration

R^2_{CV} – coefficient of determination of cross validation

R^2_{Pred} – coefficient of determination of prediction

LV – latent variables

MC – mean centring

PCA – principal component analysis

SG – Savitzky Golay

SNV – standard normal variate

SVMC – support vector machine classification

SVMR – support vector machine regression

3.1. Introduction

Eggs are one of the main natural sources of essential nutrients, such as antioxidants, amino acids, proteins, fatty acids, in addition to numerous vitamins and minerals (Bao et al.,

2020). In recent years, quail (*Nothura maculosa*) eggs have attracted a lot of attention due to their nutritional and sensory properties. Compared to chicken eggs, quail eggs have a higher nutritional value, with higher amounts of protein, antioxidants, calcium, iron and phosphorus (Bao et al., 2020; Sun et al., 2017). In Brazil, the production of quail eggs was estimated at 3,786.95 billion eggs in 2019, placing the country as the second largest producer in the world (IBGE, 2020).

Immediately after oviposition, aging affects the internal quality of egg due to physical and chemical changes in the albumen and yolk, caused by the loss of water and carbon dioxide (CO₂) through the shell, and by osmotic exchanges between the albumen and the yolk through the vitelline membrane (Giunchi et al., 2008). As consequence, there is an increase in egg acidity, liquefaction of the albumen, movement of liquids between the compartments, an increase in the air layer and distension and flaccidity of the yolk membrane (Kumbár et al., 2015). These reactions are accentuated with the increase in storage parameters such as relative humidity, time and temperature, decreasing egg quality for consumption due to deterioration of the internal quality, and higher rate transfer of CO₂ through the shell (Sari et al., 2016). Therefore, it is important to estimate the real freshness of quail eggs for consumers' safety.

Egg aging can be estimated using several parameters associated to the characteristics of egg components. The most important parameters that describe the quality of eggs during storage are albumen height, Haugh unit (HU), pH of albumen and yolk, yolk index (YI) and air chamber size (Kumbár et al., 2015). HU is the reference parameter used to estimate the freshness of chicken eggs, recognized by the Brazilian regulation of egg sanitation (Brasil, 1990). This parameter is calculated using the height of the albumen and the weight of the egg, which decreases with storage time due mainly to moisture loss. Although HU has been established for estimate chicken eggs freshness, it is used to measure the freshness of eggs from other poultry species such as quail (El-Tarabany, 2016), duck (Wang et al., 2014), partridge (Günhan & Kırıkçı, 2017) and pigeon (Chang et al., 2019). Recently, Narushin et al. (2021) proposed a new parameter called Egg Quality Index (EQI) that is proportional to the HU scale, but includes other features such as yolk height, being more representative to express chicken egg aging.

These methods are destructive, time-consuming, and require the use of toxic reagents and trained analysts. Therefore, for moving to called Industry 4.0, the food industry needs innovative and non-destructive techniques for simple, fast and economical determination of chemical and physical compounds in food. Near-infrared spectroscopy (NIRS), associated with chemometric tools, meets these requirements, finding several applications for food products,

such chemical and physical analysis, quality control, process monitoring and shelf life estimation in the food industry as an alternative to conventional analytical techniques (Grassi & Alamprese, 2018; McVey et al., 2021; Mendez et al., 2019; Walsh et al., 2020). Previous works reported the efficiency of NIRS to estimate HU and air chamber size of hen's eggs (Aboonajmi & Abbasian Najafabadi, 2014; Giunchi et al., 2008; Lin et al., 2011). On the other hand, in recent years, low cost, robust and ergonomic portable/handled NIR spectrometers were developed, allowing real-time evaluations, online applications, convenience, obtaining various simultaneous information and simple, fast and non-invasive analysis in scientific, industry and field research (Cruz-Tirado et al., 2021; He et al., 2021; Pereira et al., 2020). In this way, portable NIR spectrometers represent an attractive and viable alternative for monitoring quail eggs freshness in the entire production chain.

In the best of our knowledge, there are no reports on the application of NIR spectroscopy, bench or portable, to determine the quality attributes of quail eggs. Therefore, the objective of this work was to investigate the performance of a low-cost, portable NIR spectrometer in combination with machine learning, to estimate quail egg freshness. In addition, for the first time, we compared different freshness parameters one that best explains the aging of quail eggs.

3.2. Material and methods

3.2.1 Egg sampling

One thousand and one hundred (1100) freshly laid quail eggs (one-day old, 11.34 ± 1.03 g in weight, 3.35 ± 0.16 cm high and 2.69 ± 0.17 cm wide (**Fig. 1A**) were obtained from local farms in Campinas (São Paulo, Brazil). Unfertilized eggs were collected from the same group of quails, which were raised under the same feeding and environmental conditions. These were carefully transported from the farm to the laboratory, where they were sanitized by immersion in chlorinated water (50 ppm) at 40 ° C, for 1 minute, according to the Brazilian regulation of egg sanitation (Brasil, 1990), and analysed for the presence of cracks or damage. Subsequently, the eggs were dried at room temperature for 5 minutes, and then placed in polyethylene terephthalate (PET) trays. The samples were stored in a BOD chamber at 50 - 65% RH and 20 ° C, for a period of forty-one days. One hundred and twenty (1020) eggs were immediately analysed as a control group for the characteristics of fresh eggs. After, one hundred (100) eggs were randomly removed from storage for each day of analysis (0, 4, 7, 10, 14, 17, 21, 24, 32, 41) for spectra acquisition and determination of quality indexes (HU, EQI

and YI).

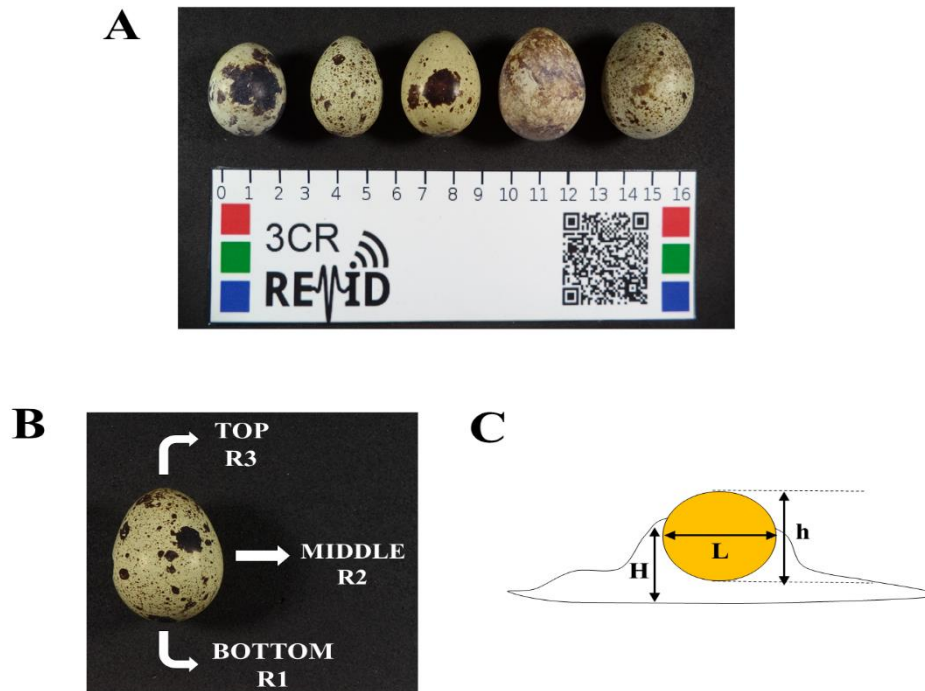


Fig. 1. Quail egg variability (A), egg region for spectral acquisition (B) and measurement region of physical parameters (C). H: height of the thick albumen (mm), h: height of the yolk (mm), L: width of the yolk (mm).

3.2.2. Spectra acquisition

The NIR spectra were acquired using a portable spectrometer (DLPR NIRscanTM Nano, Texas Instruments, USA), in absorbance mode (900 - 1700 nm) with 4 nm intervals and 20 repetitions for each spectrum, using a halogen lamp (10 W) and an InGaAs detector. Quail eggs were placed in direct contact to the spectrometer. For each egg, three detections were made, at different points: bottom region (R1), central (equatorial) region (R2) and top region (R3) (Cruz-Tirado et al., 2021) (**Fig. 1B**), in order to obtain a better representation of the samples. In this way, a total of 3060 spectra were acquired (1020 spectra $\times$ 3 points of analysis: R1, R2 and R3). The data matrices were created using the mean spectra (R1 + R2 + R3), and separately for each region (R1, R2 and R3), to compare the best region for spectral measurement and evaluate whether the addition of more measurements would improve the performance of prediction and classification models. Therefore, in total, four databases (1020 spectra for each data set) were used to calibrate and validate the models.

3.2.3. Egg quality indexes

3.2.3.1 Haugh unit (HU)

The HU is measured as a function of the height of the thick albumen and the weight of the egg. Eggs were initially weighed individually on a semi-analytical scale (± 0.0001 g), then gently broken on a flat glass surface where the height of the albumen was measured using a stainless steel analogue caliper (± 0.01 mm). Measurements were made from two points on the thick albumen (~ 10 mm from the yolk) (**Fig. 1C**). HU values were calculated using the Eq. (1) developed by Haugh (1937) for chicken eggs:

$$HU = 100x \log(H + 7.57 - 1.7xW^{0.37}) \quad \text{Eq. (1)}$$

where H is the average height of the thick albumen (millimetres) and W is the weight of the egg (grams).

3.2.3.2 Yolk index (YI)

The YI is an indicator of the spherical nature of the yolk. After measuring the height of the thick albumen (H), still on the flat glass surface the height (h) and width (L) of the yolk were measured using a stainless steel analogue micrometre (± 0.01 mm). The measurements were made from two points (**Fig. 1C**), and based on the average values obtained, the yolk index was estimated using Eq. (2) (Sharp & Powell, 1930):

$$YI (\%) = \frac{\text{Yolk height}}{\text{Yolk width}} * 100 \quad \text{Eq. (2)}$$

3.2.3.3 Egg quality index (EQI)

The egg quality index (EQI) was estimated according to the egg weight and the height of the thick albumen and yolk, according to Eq. (3) proposed by Narushin et al. (2021):

$$EQI = 100 x \text{Log} \frac{100H \sqrt{Hh}}{\sqrt{0.5W-0.5}h+(0.4W+16.5-1.3h^3)H} \quad \text{Eq. (3)}$$

where H is the average height of the thick albumen (millimeters), *h* is the average of the yolk height (millimeters) and W is the weight of the egg (grams).

3.2.4 Data analysis

The spectral data analysis was performed in four main steps: pre-treatment of the data, exploratory analysis, modelling and model performance evaluation. Data analysis was entirely performed using PLS-Toolbox from Eigenvector Research, Inc. (Manson, WA, USA) for Matlab R2019a (Mathworks, Natick, USA). The spectra were pre-processed for light scattering (derived from different data acquisition) and baseline. Therefore, after different pre-processing combinations, the best pre-processing was the combination of mean centred (mean centre - MC), the standard normal variable (Standard Normal Variate - SNV) and 1st derivative (Savitzky Golay - SG smoothing, 11 points window, 2nd polynomial order). For the construction of the prediction and classification models, the data sets were divided into two subsets: (1) calibration set (70% data) for the construction and internal cross-validation (using 10-fold cross-validation), and (2) external prediction set (30%) used to validate the predictive capacity of the models.

3.2.4.1 Exploratory analysis

To reduce the size of the data set and obtain an overview of the variation between samples, principal component analysis (PCA) was used as an exploratory analysis. This technique consists of linearly transformation of an original set of variables, initially correlated with each other, into a smaller set of non-correlated variables, called principal components (PCs), without losing significant information. The PCA was performed on pre-processed spectral data to evaluate the variation of quail egg quality during storage. In addition, the outliers (anomalous samples) were identified and eliminated using the values of $Q_{residuals}$ and T^2 Hotellings.

3.2.4.2. Prediction of egg quality parameters

Partial least squares regression (PLSR) and support vector machines regression (SVMR) were built, based on pre-processed spectra, to predict egg quality parameters (HU, YI and EQI). The choice of the number of latent variables (LVs) that should be included in the PLSR models was made based on the smallest mean square error of the cross-validation (RMSECV), with a maximum of 12 LVs. PLSR is the most popular regression model for building NIR prediction model, which is based on the linear regression relationship between NIR data and dependent variables. SVMR is a supervised and nonparametric kernel-based learning algorithm (Vapnik, 2013), which uses training vector to defines the best sparse deterministic regression relationship between the NIR data and the response values (Grelet et

al., 2020). In contrast to PLSR, SVMR is able to model nonlinear relations, and it also can efficiently work with high dimensional input vectors (Wang et al., 2014). In this study, SVMR was used after compression of the data by PLS using 10 latent variables. The SVMR models were developed using a radial kernel base function (RBF), and their optimization based on the parameters c (cost) and ϵ (epsilon) with a grid search, based on the smallest RMSECV of a 10-fold cross validation.

The reliability and robustness of the regression models were evaluated by the calculation of the root of the mean square error of calibration, cross-validation and prediction (RMSEC, RMSECV and RMSEP), of the error range rate (RER), provided from the ratio between the calibration interval (interval) and the RMSEP (Eq. (4)), and the residual prediction deviation (RPD), as the ratio between the standard deviation of the calibration set (DPcal) and the average values of the RMSEP (Eq. (5)). RER values > 10 are indicative of models with a good estimate (Páscoa et al., 2013). RPD values < 1.5 indicate that the model cannot be used, $RPD = 1.5 - 2.0$ states that the model can distinguish between high and low values, $RPD = 2.0 - 2.5$ evolution a model with approximate prediction capacity, $RPD = 2.5 - 3.0$ indicates a good model for prediction, and $RPD > 3$ requiring excellent predictive capacity of the model (Saeys et al., 2005).

$$RER = \frac{range}{RMSEP} \quad \text{Eq. (4)}$$

$$RPD = \frac{DPcal}{RMSEP} \quad \text{Eq. (5)}$$

3.2.4.3 Egg freshness classification

Partial least squares discriminant analysis (PLSDA) and the support vector machine classification (SVMC) were explored in order to assess the efficiency of the portable NIR in classifying quail eggs according to their freshness (*fresh* and *stale*), based on HU and EQI values. The choice of the number of latent variables (LVs) that should be included in the PLSDA models was made based on the smallest mean square error of the cross-validation (RMSECV), with a maximum of 12 LVs. The classification by SVMC obtains an optimal hyperplane for which the margin of separation between the different classes is maximum. In this study, SVMC was compacted using 10 PCs. The SVMC models were developed using a radial kernel base function (RBF), and their optimization based on the parameters c (cost) and ϵ (epsilon) with a grid search, based on the smallest RMSECV of a 10-fold cross validation.

The performance of the classification models was evaluated by the sensitivity, specificity, precision and error rates for the calibration/cross-validation and prediction sets. Sensitivity (Eq. (6)) and specificity (Eq. (7)) can be described as the model's ability to correctly classify positive and negative samples, respectively. The precision (Eq. (8)) and error (Eq. (9)) rates are statistical parameters that provide a global value to measure the model's performance and error. These are calculated using the ratio between the number of samples correctly classified and misclassified, regardless of the class, and the total number of samples, respectively.

$$\text{Sensitivity (\%)} = \frac{TP}{(TP+FN)} \times 100 \quad \text{Eq. (6)}$$

$$\text{Specificity (\%)} = \frac{TN}{(TN+FP)} \times 100 \quad \text{Eq. (7)}$$

$$\text{Accuracy (\%)} = \frac{(TP+TN)}{Total} \times 100 \quad \text{Eq. (8)}$$

$$\text{Error (\%)} = \frac{(FP+FN)}{Total} \times 100 \quad \text{Eq. (9)}$$

where TP = true positive, TN = true negative, FP = false positive and FN = false negative.

3.2.5 Statistical analysis

The analysis of variance (ANOVA) followed by the Tukey test ($p < 0.05$) were performed to compare the averages of quality parameters of quail eggs during storage. The software used to perform the statistical calculations was Statistica version 8.0 (Statsoft, Oklahoma, USA).

3.3 Results and discussion

3.3.1 Freshness of eggs during storage

The evolution of quail eggs freshness was calculated considering quality parameters such as the Haugh unit (HU), yolk index (YI) and egg quality index (EQI), as well as experimental observations (**Fig. 2**). Due to the lack of previous references on the shelf life of quail eggs, experimental observations, such as observation of the size of the air layer and the density reduction caused by weight loss, were essential to verify whether the eggs were fresh or stale.

Fig. 2 shows that the HU decreases significantly (Tukey's test, $p < 0.05$) only after 10

days of storage, showing a decrease of only 21.55% at the end of the storage period. (Akpınar et al., 2015) observed a decrease of 13.03% for quail eggs stored at 25 ° C for 35 days. (Kumbár et al., 2015) also reported a small change of 4.38% in HU values, for quail eggs stored for 42 days at 4 ° C. During the storage of eggs, the pH of the albumen increases due to CO₂ losses through the shell, which cause the hydrolysis of the amino acid chains, which release water that was linked to large protein molecules, mainly ovomucine, resulting in liquefaction of thick albumen (de Moraes Garcia et al., 2015). Therefore, if we use the HU scale as established for chicken eggs, all quail eggs would be classified as fresh, which is not consistent with experimental observations (**Fig. 2**). As for the yolk index (YI), it is possible to observe a significant decrease (Tukey's test, $p < 0.05$) during storage. This is due to the increased permeability and weakening of the vitelline membrane, facilitating the displacement of excess water from the albumen to the yolk, making it flabby and flattened (Marinho, 2011), which is visible in digital images of boiled eggs (**Fig. 2**). Similarly, the egg quality index (EQI) decreased with storage time (**Fig. 2**). Due to the fact that EQI is a function of both the height of the albumen and yolk, this decreasing was probably caused by the degradation of the yolk, instead of changes in the albumen, which is supported both by the experimental observations (**Fig. 2**) and by the evolution of the HU and YI parameters. Therefore, it is possible to conclude that the storage period is a crucial factor affecting quail eggs quality. However, it should be noted that YI and EQI were more significantly influenced by the storage time, when compared to HU (**Fig. 2**), reinforcing the idea that they can more accurately describe the freshness of quail eggs stored at 20 ° C.

On the other hand, EQI has a high correlation with HU, which means that the values defined in the HU scale to classify the freshness of chicken eggs are practically the same for EQI (Class AA: $EQI \geq 72$; Class A: $EQI = 71 - 60$; Class B: $EQI < 60$; or $EQI \geq 70$ for fresh eggs and $EQI < 70$ for stale eggs). Therefore, analysing **Fig. 2**, we can define that eggs stored until the 14th day can be considered fresh. Previous work has shown that chicken eggs stored at 20 ° C remained fresh until the 14th day (Cruz-Tirado et al., 2021; Liu et al., 2020). In addition, still based on the correlation between EQI and HU, it is possible (maybe) to redefine the HU scale specifically for quail eggs, with $HU \geq 80$ for fresh eggs (related to day 14) and $HU < 80$ for stale eggs.

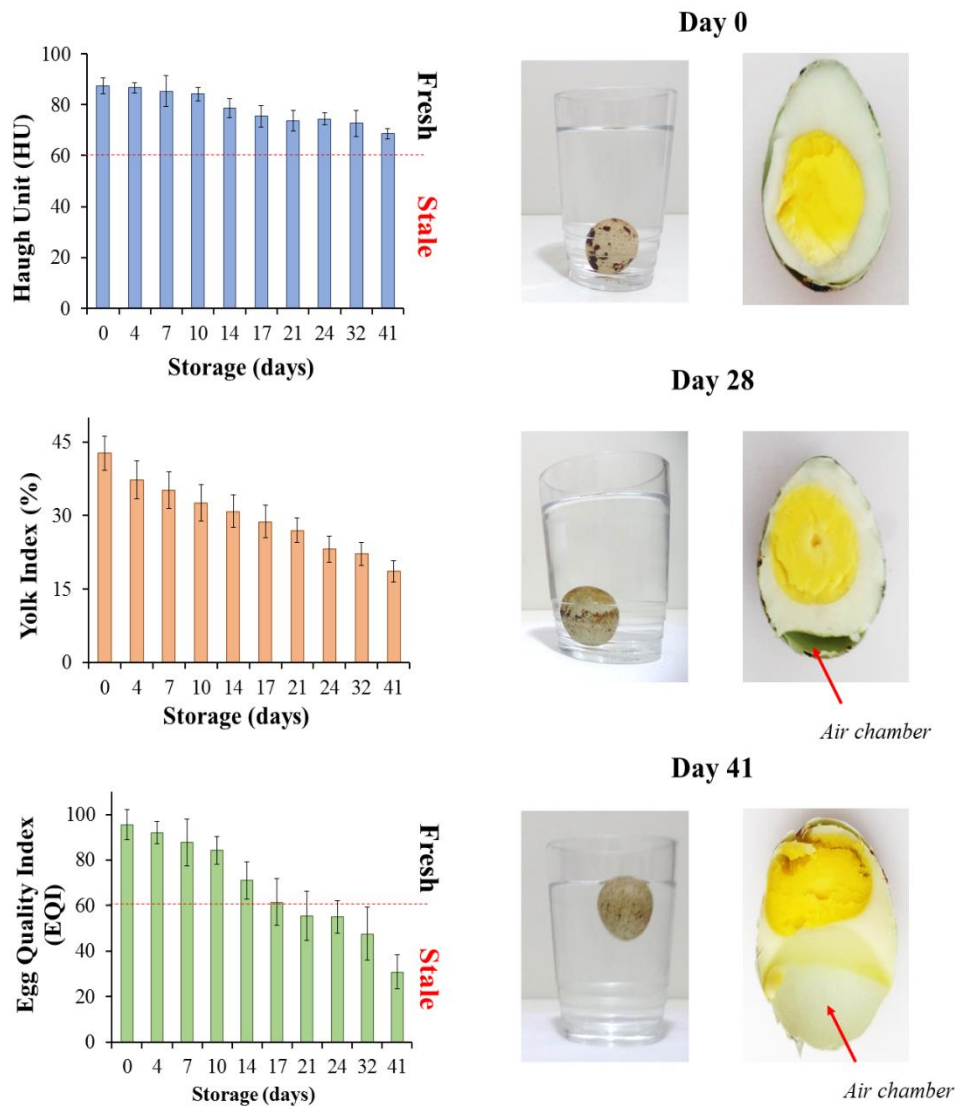


Fig. 2. Average value of the Haugh Unit (HU), Yolk Index (YI) and Egg Quality Index (EQI) and experimental observations of quail eggs stored for 41 days at 20 °C.

3.3.2. Spectral and exploratory analysis

Fig. 3 shows the raw spectra (**Fig. 3A**), pre-processed spectra by SNV + 1st S-G derivative (**Fig. 3B**), PCA analysis based on storage time (**Fig. 3C**), PCA analysis based on freshness grading (*fresh* or *stale*) (**Fig. 3D**) and loadings plot (**Fig. 3E**). PCA analysis was carried out in two forms: (1) to observe the spectra variations of quail eggs samples during storage, which include nine classes (or days), and (2) freshness grading (*fresh* and *stale*).

Fig. 3A and **3B** show the graphs of the raw and pre-processed mean spectra. The difference in the baseline between fresh and stale eggs observed in the raw spectra (**Fig. 3A**), corresponds to the change in light scattering properties, probably associated with the air

chamber and the degradation of the inner film of quail eggs. The regions of 938 nm and the peak around 1334 nm, are associated with the C-H elongation, characteristic of saturated fatty acids (fat) present in the yolk (Hourant et al., 2000). The absorption band at 1188 nm have been described as the second overtone of C-H absorption of pure fatty acids containing *cis* double bonds (Sato et al., 1991), which could be associated with the higher content of oleic acid in quail egg yolk (Arantes da Silva et al., 2009; Choi et al., 2001). The peak observed at 989 and 1400 nm is associated with the O-H stretching and elongation, respectively, referring to water. The difference in absorbance in this region may be linked to the loss of moisture through the pores of the shell, as well as the structural change of proteins in quail eggs during storage (Cruz-Tirado et al., 2021). The peaks around 1334 and 1643 nm are related to the second aromatic C-H elongation overtone, mainly related to CH₂ and CH₃ (Kaufmann et al., 2019). The increase in absorbance for stale quail eggs is probably associated with the release of carbon dioxide by the shell (Jiang et al., 2018).

Fig. 3C shows the score plot for spectral data according to storage day, with PC1 explains 83.74% and PC2 explains 6.59%. It is possible to observe as samples were moving from positive to negative scores on PC1 with the increasing of storage day. In addition, the freshest samples (Day 0) and oldest sample (Day 41) were characterized for negative scores on PC2. The egg samples from Day 14 is overlapped in the centre of PCA score plot, indicating a transition phase from fresh (0, 4, 7 days) to stale (17, 21, 24, 32, 41) grading. Based on the physicochemical analyses (**Fig. 2**), it is possible to consider that the samples with more than 14 days of storage are stale. Therefore, **Fig. 3D** was plotted using threshold of 14 days for clustering fresh (≤ 14 days) and stale (> 14 days). Although not clearly observed, the majority of fresh eggs are located on positive PC1 scores, while stale eggs are located in negative PC1 scores. The loadings plot in **Fig. 3E** shows that most of the peaks and valleys that contributed for sample clustering are those associated with water, fat and aromatic compounds. This indicates that the spectral information collected with the portable NIR is strongly associated with the variations in the composition of the quail eggs during storage, and can be used for a screening of eggs according to their freshness.

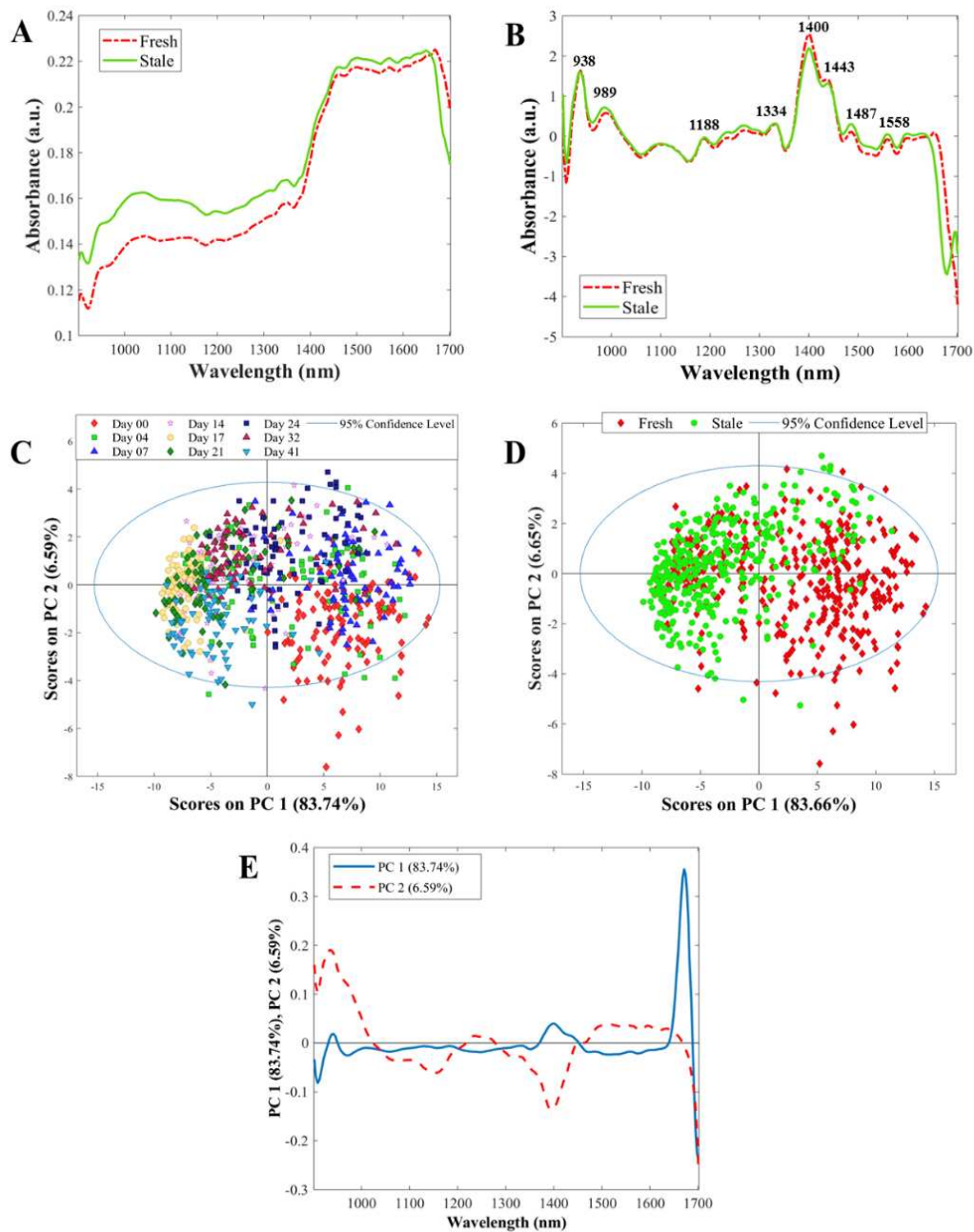


Fig. 3. Raw NIR spectra (A), pre-processed spectra (SNV + 1st derivative – SG) (B), score plot based on storage time (C), score plot based on freshness grading (*fresh* or *stale*) (fresh and stale) (D) and loading plot (E).

3.3.3. Prediction of egg quality indexes

The prediction performance of egg quality indexes (HU, YI and EQI) using PLSR and SVMR models are shown in **Fig. 4** and **Fig. 5**, respectively. In this study, the performance of the models was preferably evaluated based on error (RMSECV and RMSEP) and related variables (RER and RPD), since R^2 is dependent on the range of the data set (Davies & Fearn,

2006). The best models are considered those that have lower values of RMSECV and RMSEP and higher values of RPD and RER, as these measurements are made to quantitatively verify the precision of the predictive capacity of the chemometric models (Conzen, 2006). In general, the SVMR models showed better performance than PLSR models, which leads us to suggest that perhaps the relationship between the NIR spectra and the parameters HU, YI and EQI tends to be non-linear instead of linear. Thus, linear methods may not completely show the relationship between the NIR spectra and the freshness of quail eggs. The best performance of PLSR was achieved for data set from mean spectra of three measurement points (R1 + R2 + R3) for all quality indexes (**Table 1**). In contrast, the models built using the SVMR algorithm showed the best performance in the R3, R2 and mean spectra for HU, IG and EQI predictions, respectively (**Table 1**). Based on the RPD and RER, for the PLSR only the models for estimating the YI (RPD = 2.00 - 2.13 and RER = 10.25 - 10.96), and for the SVMR only the models to estimate YI (RPD = 2.07 - 2.31 and RER = 10.63 - 11.87) and EQI (RPD = 2.14 - 2.35 and RER = 9.19 - 10.08) indicate a good correlation between the values predicted by the model and the real (experimental) values. **Fig. 3** and **Fig. 4** show the best PLSR and SVMR models, separately (R1; R2; R3) and based on the mean (R1 + R2 + R3), respectively.

Table 1. Performance of the PLSR and SVMR models for prediction of the Haugh Unit (HU), Yolk Index (YI) and Egg Quality Index (EQI) of quail eggs.

Model	Index	Spectra origin	RMSEC	RMSECV	RMSEP	R^2_{CAL}	R^2_{CV}	R^2_{PRED}	RPD	RER
PLSR	HU	R1	3.90	4.15	4.12	0.68	0.63	0.65	1.70	7.11
		R2	4.02	4.23	4.18	0.66	0.63	0.64	1.67	7.01
		R3	3.73	3.99	4.24	0.71	0.67	0.65	1.65	6.92
		Mean	3.70	3.91	3.90	0.71	0.67	0.69	1.79	7.50
	YI	R1	3.71	3.97	3.89	0.76	0.72	0.76	2.00	10.25
		R2	3.46	3.80	3.79	0.79	0.75	0.77	2.05	10.53
		R3	3.58	3.82	3.73	0.78	0.75	0.77	2.08	10.67
		Mean	3.45	3.65	3.64	0.80	0.77	0.79	2.13	10.96
	EQI	R1	10.94	11.68	11.31	0.72	0.68	0.70	1.92	8.25
		R2	11.23	11.93	11.41	0.71	0.67	0.70	1.91	8.18
		R3	10.49	11.17	11.95	0.75	0.72	0.70	1.82	7.81
		Mean	10.54	11.16	11.02	0.75	0.72	0.72	1.97	8.47

SVMR	HU	R1	2.36	3.80	3.76	0.89	0.70	0.71	1.86	7.79
		R2	2.66	3.65	3.70	0.85	0.72	0.73	1.89	7.92
		R3	3.02	3.67	3.69	0.81	0.72	0.72	1.90	7.94
		Mean	3.12	3.74	3.71	0.79	0.70	0.73	1.88	7.90
	YI	R1	3.12	3.75	3.75	0.83	0.76	0.78	2.07	10.63
		R2	2.59	3.57	3.36	0.89	0.78	0.81	2.31	11.87
		R3	2.95	3.66	3.60	0.85	0.77	0.79	2.15	11.06
		Mean	3.11	3.52	3.37	0.83	0.79	0.81	2.30	11.82
	EQI	R1	6.40	10.31	9.93	0.91	0.76	0.78	2.19	9.40
		R2	7.11	10.17	10.03	0.89	0.77	0.79	2.17	9.30
		R3	7.26	10.12	10.15	0.88	0.77	0.77	2.14	9.19
		Mean	6.81	9.42	9.26	0.90	0.80	0.82	2.35	10.08

HU regression models showed the lowest R^2_p values (0.63 – 0.73) and the higher RMSEP (3.69 – 4.12) for both PLSR and SVMR model (**Fig. 1**). This was probably because the R^2 is dependent on the rank of the reference values and HU had the lowest reference range. In addition, the best RPD value reached 1.90 using SVMR model based on spectra from R3 region. Based on RPD scale, the model can be considered for screening between higher and lower values (Saeys et al., 2005). During storage, R3 is located in the egg region of greatest contact with the albumen (top), making the light from the NIR sensor, passing through the pores of the egg shell, stay in direct contact with the albumen during the period evaluated. In contrast, regions R2 and R1 are the region where the albumen liquefies during storage causing the yolk to approach the shell (R2) and to increase the air chamber (R1) (**Fig. 2**).

For YI, the SVM and PLS models presented similar performance with RPD ranging from 2.0 - 2.5 and RER > 10, which suggests a model with reasonable prediction capacity. However, the best PLSR model (RPD = 2.13) required spectral information from the three regions, while the best SVMR model (RPD = 2.30) required only spectra from R2 region, optimizing the time to collect and process spectral information. This may be related to the fact that during storage the yolk approaches the shell (**Fig. 2**), allowing a better interaction with the NIR, and consequently obtaining more representative spectral information in R2. The egg spectral profile (**Fig. 3B**) shows characteristic bands associated with the lipid compounds present in the yolk.

Regarding the EQI, the SVMR models (RPD = 2.35 and RER = 10.08) showed a better

predictive capacity, when compared to the PLSR models ($RPD = 1.97$ and $RER = 8.47$). These results indicate that the SVMR model has a reasonable prediction capacity for EQI values. The best models, for both algorithms, were obtained using the mean spectra of the three regions. A better performance for the EQI regression models, compared to the HU models, may be related to the fact that the spectral information of the NIR presents representative absorption peaks associated with the yolk components (Fig. 3B).

As stated by (Dong et al., 2019), the eggshell thicknesses at the bottom (R1), middle (R2) and top (R3) are different, so that each region would have an influence on the spectra. In addition, the composition of egg white and yolk, and the composition of egg cuticle, change with storage time (Samiullah & Roberts, 2014). In R2 region, when the light falls on the pores of the shell and comes into contact with the cuticle, which in turn is in contact with albumen, the spectrum can better detect the freshness of the egg. On the other hand, the air chamber is located at the bottom of the egg (R1), with the diameter and height increasing during the storage. Therefore, they have a high correlation with the freshness of the egg (Liu et al., 2020). Cruz-Tirado et al. (2021) also reported that the better predictive capacity for hen's eggs was reached for HU regression models based on mean spectra of three measurements using SVMR algorithm.

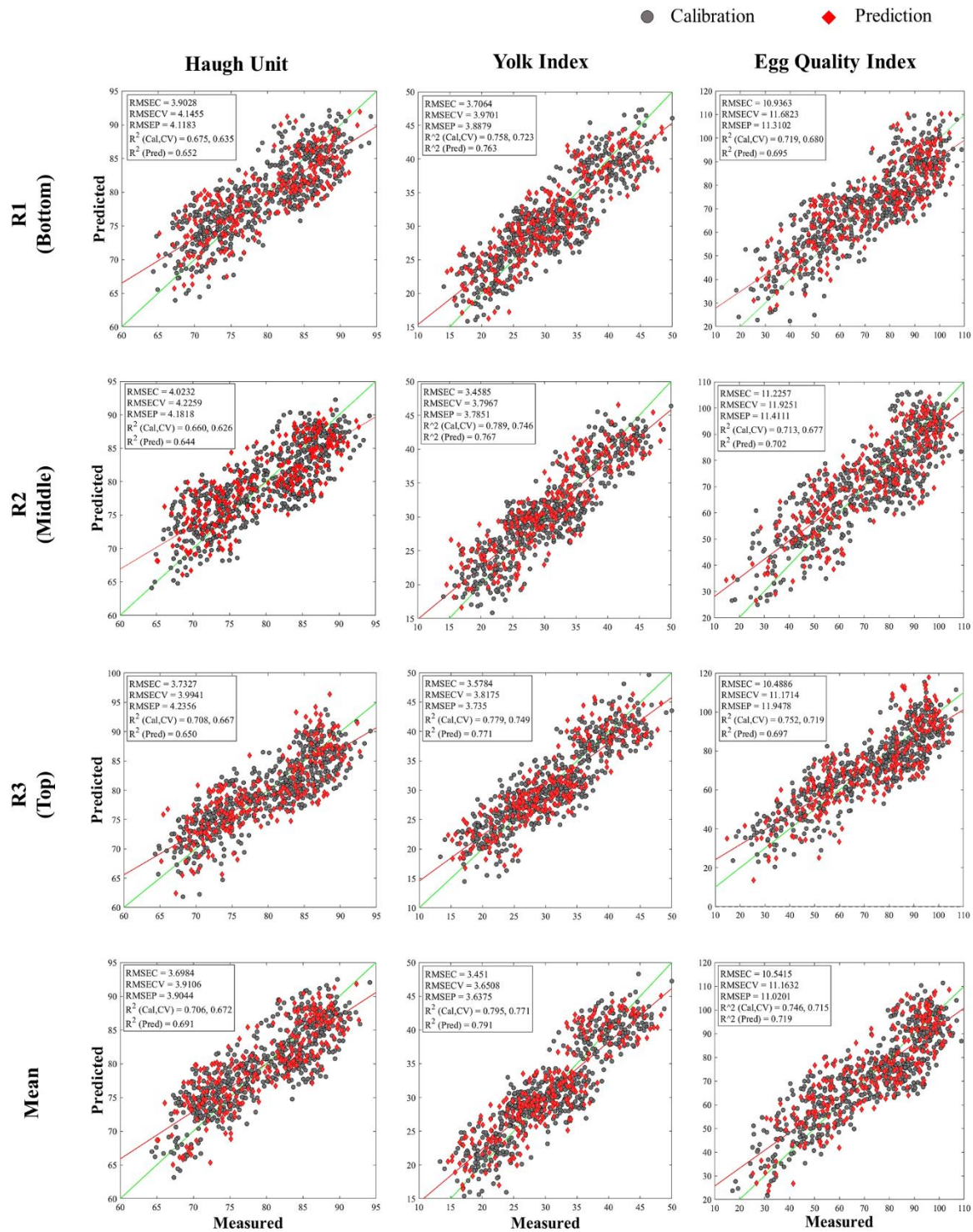


Fig. 4. PLSR models to predict HU, IG and EQI in the three different measurement points (R1: bottom, R2: middle, R3: top). Grey circle are calibration samples, and red diamonds are validation samples.

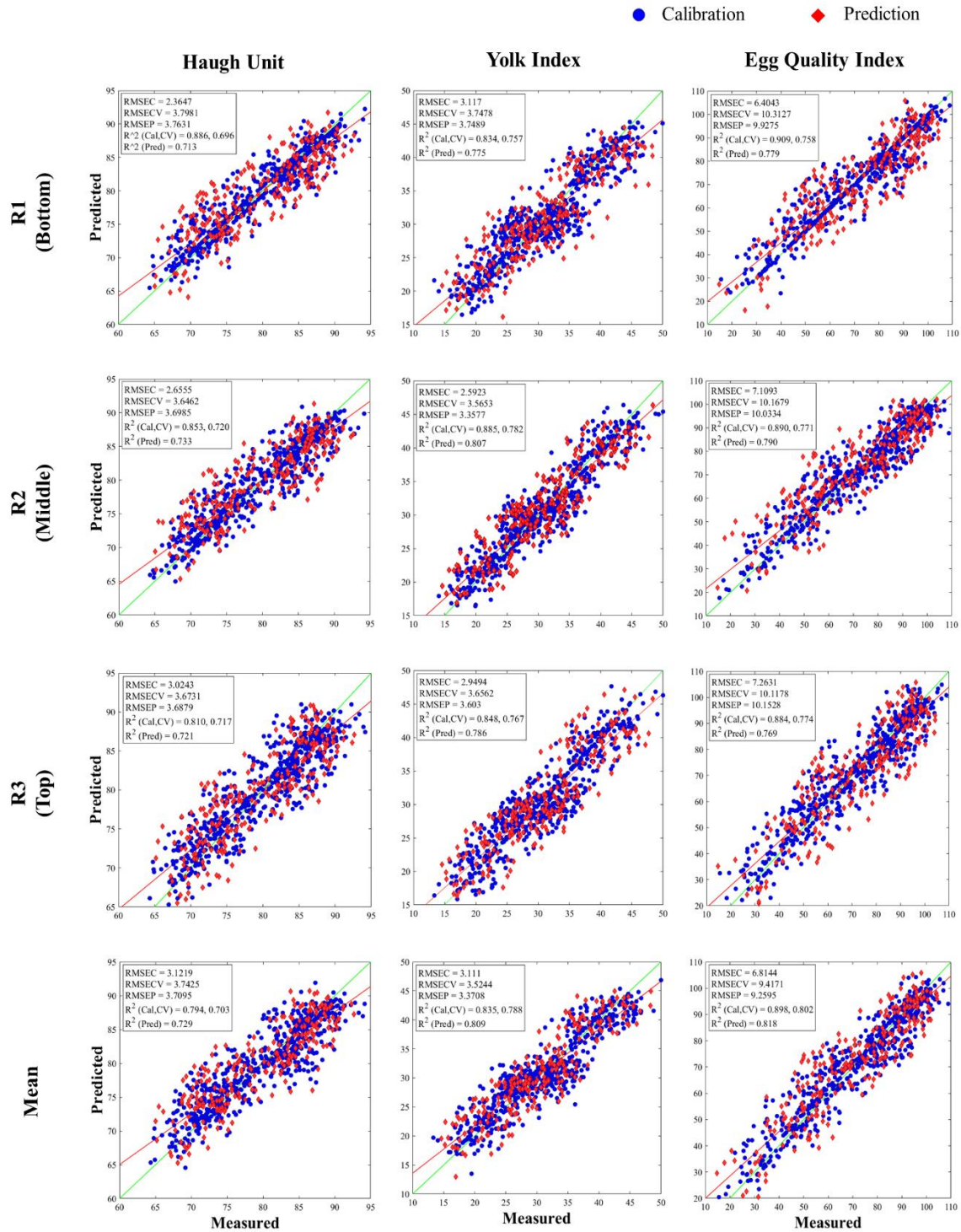


Fig. 5. SVMR models to predict HU, IG and EQI in the three different measurement points (R1: bottom, R2: middle, R3: top).

3.3.4. Egg freshness grading

Table 2 shows the performance of the PLSDA and SVMC classification models

according to the error rates, sensitivity, specificity and precision for each spectral region separately (R1; R2; R3) and for the mean spectra (R1 + R2 + R3). The confusion matrices for classification results of the PLSDA and SVMC models into “*fresh*” ($HU \geq 80$; $EQI \geq 70$) and “*stale*” ($HU < 80$; $EQI < 70$) are shown in **Fig. 6** and **Fig. 7**, respectively.

The PLSDA and SVMC models for classifying quail eggs into “*fresh*” and “*stale*”, defined by HU and EQI values, showed similar performance (**Table 2**). This indicates that the restructuring of the HU scale (based on the correlation with EQI, (Narushin et al., 2021) proposed in this study can properly express quail egg freshness, which can be used for classification purposes. The best detection region for the construction of the PLSDA and SVMC models was observed in the mean spectra. This behavior is probably related to the fact that the medium spectra are more representative to describe the changes in the characteristics of all the components of the quail eggs (air chamber, albumen and yolk). The best PLSDA and SVMC model was able to correctly classify 110/133 (82.7%) and 107/129 (82.9%) ‘fresh’ and 101/114 (87.8%) and 101/115 (85, 4%) ‘old’ for HU, while 143/160 (89.4%) and 150/174 (86.2%) in ‘new’ and 71/88 (80.7%) and 60/70 (85, 7%) ‘old’ were classified for EQI (**Fig. 6 and 7**), respectively. These models achieved rates of sensitivity and specificity greater than 80.0 % and error rates less than 20.0 % for both classes (Table 2). Most of the classes classified incorrectly (**Fig. 6 and 7**) belong to the transition group (days 14 and 17), as these have similar HU and EQI values, but belong to different classes. Therefore, the PLSDA and SVMC models can be used to identify and discriminate quail fresh eggs from stale eggs, which is very useful for the egg products industry.

Table 2. Performance of PLSDA and SVMC models for classification of quail eggs according to their freshness using Haugh Unit (HU) and Egg Quality Index (EQI).

Model Index	Spectra origin		Error (%)	SEN (%)	SPE (%)	ACC (%)
PLSDA	HU	Cal	12.1	87.8	88.1	87.9
		R1 CV	14.5	85.8	85.3	85.5
		Pred	15.0	89.6	80.5	85.0
	HU	Cal	10.7	88.1	90.5	89.3
		R2 CV	13.8	85.5	87.0	86.2
		Pred	14.8	90.2	80.3	85.3
	R3	Cal	8.5	91.1	91.8	91.5

SVMC	EQI		CV	12.3	87.5	87.9	87.7
			Pred	17.7	81.3	83.2	82.3
		Mean	Cal	7.7	92.8	91.8	67.5
			CV	10.1	90.2	89.7	89.9
			Pred	14.6	89.4	81.5	85.4
		R1	Cal	11.7	90.6	86.0	88.3
			CV	15.1	87.9	81.9	84.9
			Pred	18.1	92.0	71.8	81.5
		R2	Cal	11.9	92.6	83.6	88.1
			CV	14.8	89.4	81.0	85.2
			Pred	20.5	92.4	66.7	79.3
		R3	Cal	8.2	93.7	89.9	91.8
	CV		10.8	92.2	86.2	89.1	
	Pred		16.8	86.8	79.5	83.2	
	HU	Mean	Cal	8.3	93.3	90.1	91.8
			CV	10.8	92.0	86.4	89.1
			Pred	15.0	89.4	80.7	85.0
		R1	Cal	11.0	88.3	89.7	90.6
			CV	17.4	81.5	83.8	82.6
			Pred	17.5	82.4	82.7	82.5
		R2	Cal	12.5	89.6	85.5	89.8
			CV	15.5	85.7	83.3	84.5
			Pred	17.6	86.7	78.2	82.5
		R3	Cal	9.6	91.9	89.0	92.8
CV			15.5	85.7	83.3	84.5	
Pred			15.7	83.1	85.6	84.3	
Mean	Cal	8.4	93.8	89.3	93.8		
	CV	13.6	89.0	83.7	86.4		
	Pred	14.7	88.4	82.1	85.3		
	R1	Cal	17.8	95.0	69.4	85.1	
		CV	18.9	94.0	68.1	81.1	
		Pred	20.2	90.0	69.5	79.8	

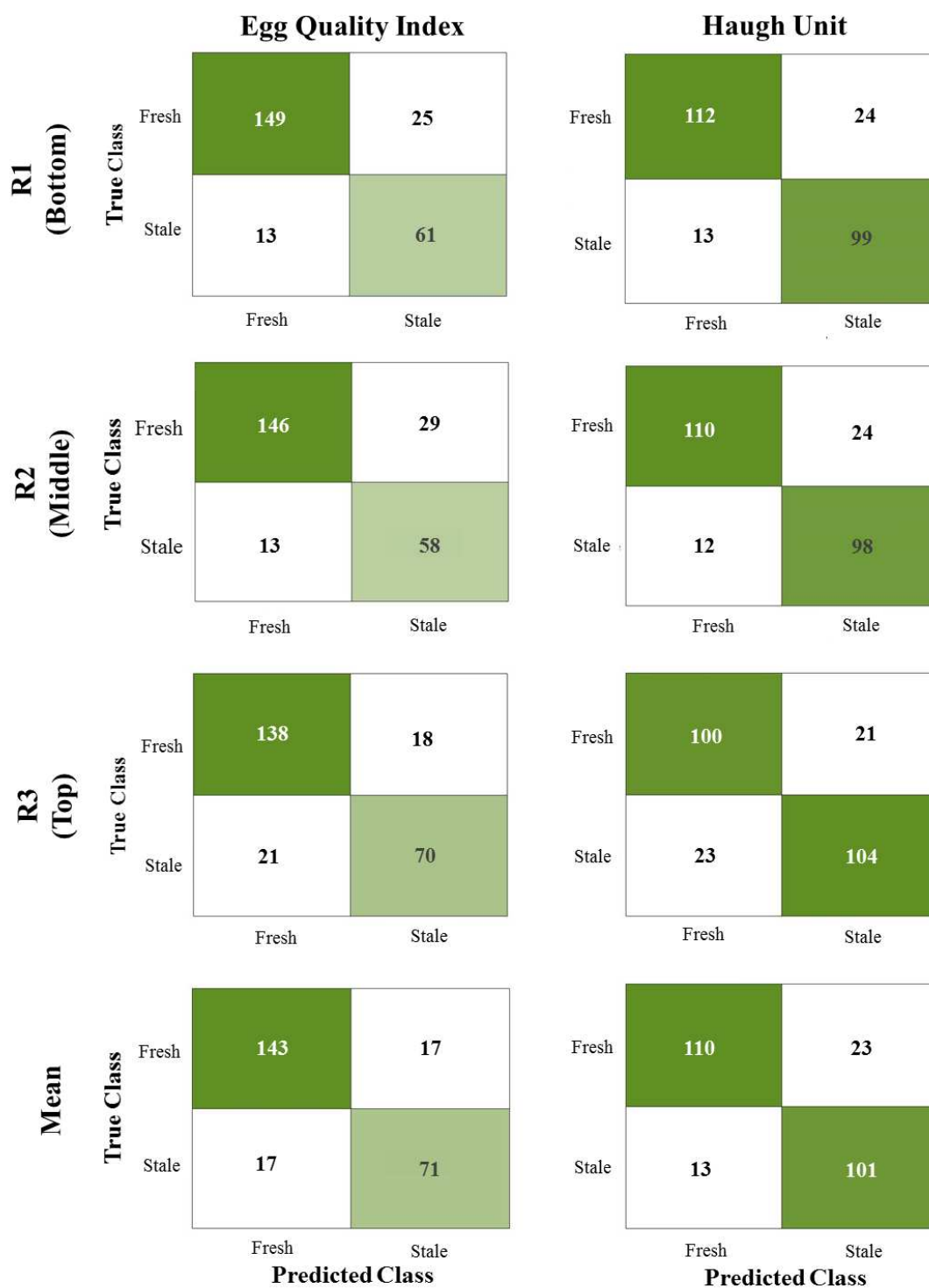


Fig. 6. PLSDA models for classification of quail eggs according to their freshness: confusion matrix for Haugh Unit (HU) and Egg Quality Index (EQI).

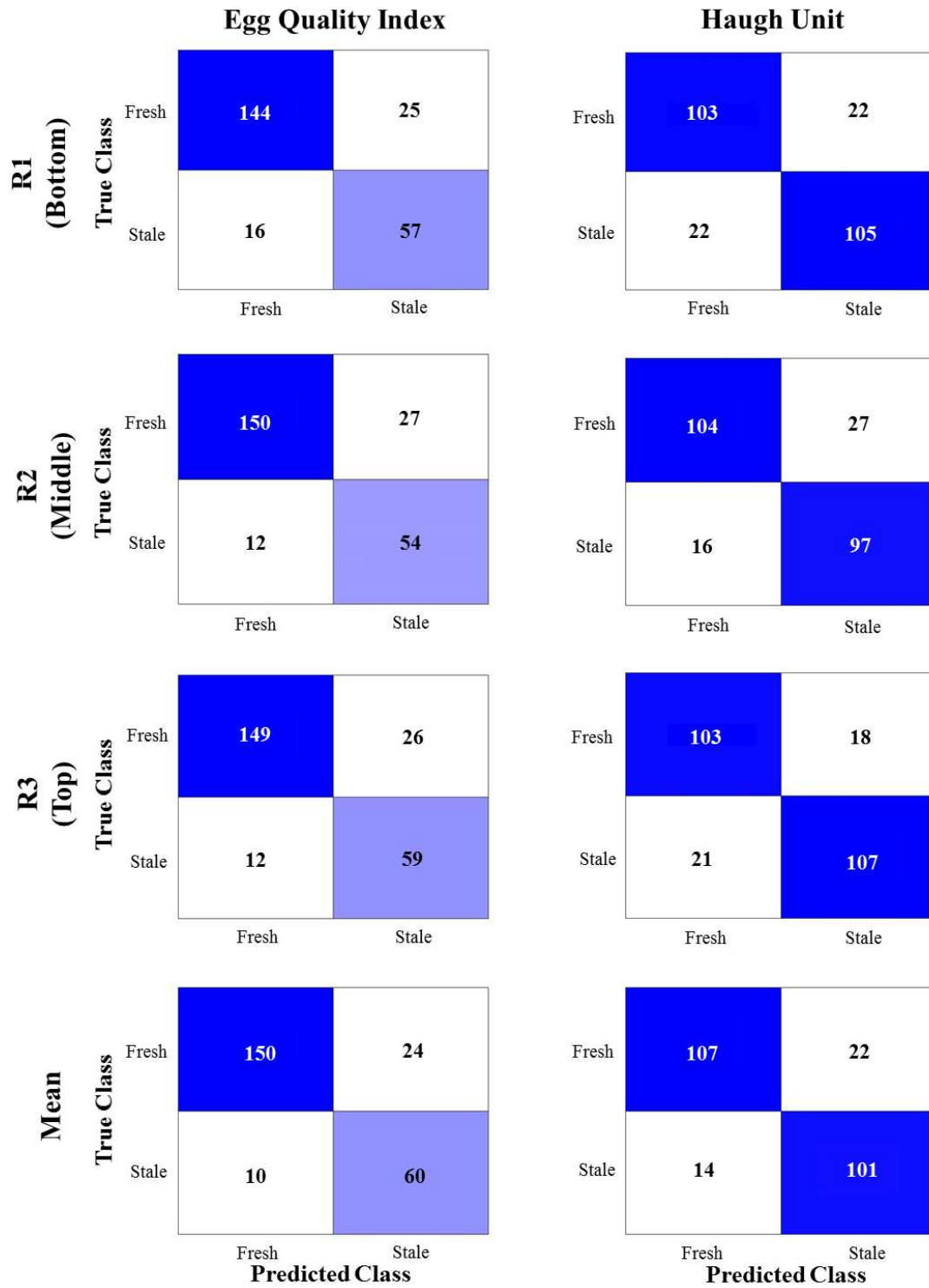


Fig. 7. SVMC models for classifying quail eggs according to their freshness: confusion matrix for Haugh Unit (HU) and Egg Quality Index (EQI).

4. Conclusion

We demonstrate that the combination of NIR spectroscopy with chemometric tools provides a fast, non-invasive and high-performance system for monitoring internal quality parameters and assessing the freshness of quail eggs during storage. EQI is a better parameter to represent quail egg freshness when compared to HU. Hence, we suggest the utilization of EQI to determine quail egg freshness, and also that HU scale for quail eggs to be restructured as $HU \geq 80$ for fresh eggs and $HU < 80$ for old eggs. The regression models with greater predictive capacity and robustness were obtained for YI and EQI, using only one spectral region (R2) when developed through SVMR. In contrast, for PLSR models it was necessary to use spectral information of three points ($R1 + R2 + R3$). The best PLSDA and SVMC models for HU and EQI correctly classified more than 80% of the samples, due to the reconfiguration of the HU scale proposed in this study. Finally, it can be concluded that the physical and chemical properties of quail eggs are closely related to the storage period, and that NIR spectroscopy is a promising alternative for assessing the freshness of quail eggs, replacing conventional destructive methods.

Acknowledgment

This research was subsidized by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code **001** and São Paulo Research Foundation (FAPESP) (project number 2020 / 09198-1, 2018 / 02500-4, 2015 / 24351-2). Yasmin Lima Brasil thanks the FAPESP grant, grant nº 2019 / 11896-1. We are also grateful to our academic colleagues for their contributions.

Conflict of interest

The authors declare no competing financial interest.

References

- Aboonajmi, M., & Abbasian Najafabadi, T. (2014). Prediction of Poultry Egg Freshness Using Vis-Nir Spectroscopy with Maximum Likelihood Method. *International Journal of Food Properties*, 17(10), 2166–2176. <https://doi.org/10.1080/10942912.2013.784330>
- Akpınar, G. C., Canogullari, S., Baylan, M., Alasahan, S., & Aygun, A. (2015). The use of propolis extract for the storage of quail eggs. *Journal of Applied Poultry Research*, 24(4), 427–435. [https://doi.org/https://doi.org/10.3382/japr/pfv043](https://doi.org/10.3382/japr/pfv043)
- Arantes da Silva, W., Naiverti Elias, A. H., Aricetti, J. A., Sakamoto, M. I., Murakami,

A. E., Marques Gomes, S. T., Visentainer, J. V., Evelázio de Souza, N., & Matsushita, M. (2009). Quail egg yolk (*Coturnix coturnix japonica*) enriched with omega-3 fatty acids. *LWT - Food Science and Technology*, 42(2), 660–663. <https://doi.org/https://doi.org/10.1016/j.lwt.2008.08.005>

Bao, Z., Kang, D., Li, C., Zhang, F., & Lin, S. (2020). Effect of salting on the water migration, physicochemical and textural characteristics, and microstructure of quail eggs. *LWT*, 132(1), 109847. <https://doi.org/10.1016/j.lwt.2020.109847>

BRASIL. (1990). *Ministério da Agricultura, Pecuária e Abastecimento. Portaria n° 1, de 21 de fevereiro de.*

Chang, L., Zhang, R., Fu, S., Mu, C., Tang, Q., & Bu, Z. (2019). Effects of Different Dietary Calcium Levels on the Performance, Egg Quality, and Albumen Transparency of Laying Pigeons. In *Animals* (Vol. 9, Issue 3). <https://doi.org/10.3390/ani9030110>

Choi, S. H., Song, K. T., & Oh, H. R. (2001). Cholesterol Contents and Fatty Acid Composition of Chukar, Pheasant, Guinea Fowl and Quail Egg Yolk. *Asian-Australas J Anim Sci*, 14(6), 831–836. <https://doi.org/10.5713/ajas.2001.831>

Conzen, J. P. (2006). Multivariate Calibration: A practical guide for developing methods in the quantitative analytical chemistry. *Bruker Optik GmbH: Ettlingen, Germany*.

Cruz-Tirado, J. P., Lucimar da Silva Medeiros, M., & Barbin, D. F. (2021). On-line monitoring of egg freshness using a portable NIR spectrometer in tandem with machine learning. *Journal of Food Engineering*, 306, 110643. <https://doi.org/https://doi.org/10.1016/j.jfoodeng.2021.110643>

Davies, A. M. C., & Fearn, T. (2006). Back to basics: calibration statistics. *Spectroscopy Europe*, 18(2), 31–32.

de Moraes Garcia, E. R., de Ávila, L. R., da Cruz, F. K., de Pedro Souza, R. P., Batista, N. R., Feliciano, W. B., & Arguelo, N. N. (2015). Qualidade de ovos de codornas japonesas: efeito da idade da ave, temperatura de conservação e período de armazenamento. *Arquivos de Ciências Veterinárias e Zoologia Da UNIPAR*, 18(4).

Dong, X., Dong, J., Li, Y., Xu, H., & Tang, X. (2019). Maintaining the predictive abilities of egg freshness models on new variety based on VIS-NIR spectroscopy technique. *Computers and Electronics in Agriculture*, 156(October 2018), 669–676. <https://doi.org/10.1016/j.compag.2018.12.012>

El-Tarabany, M. S. (2016). Effect of thermal stress on fertility and egg quality of Japanese quail. *Journal of Thermal Biology*, 61, 38–43. <https://doi.org/10.1016/j.jtherbio.2016.08.004>

Giunchi, A., Berardinelli, A., Ragni, L., Fabbri, A., & Silaghi, F. A. (2008). Non-destructive freshness assessment of shell eggs using FT-NIR spectroscopy. *Journal of Food Engineering*, 89(2), 142–148. <https://doi.org/10.1016/j.jfoodeng.2008.04.013>

Grassi, S., & Alamprese, C. (2018). Advances in NIR spectroscopy applied to process analytical technology in food industries. *Current Opinion in Food Science*, 22, 17–21. <https://doi.org/https://doi.org/10.1016/j.cofs.2017.12.008>

Grelet, C., Froidmont, E., Foldager, L., Salavati, M., Hostens, M., Ferris, C. P., Ingvarsten, K. L., Crowe, M. A., Sorensen, M. T., Fernandez Pierna, J. A., Vanlierde, A., Gengler, N., & Dehareng, F. (2020). Potential of milk mid-infrared spectra to predict nitrogen use efficiency of individual dairy cows in early lactation. *Journal of Dairy Science*, 103(5), 4435–4445. <https://doi.org/https://doi.org/10.3168/jds.2019-17910>

Günhan, Ş., & Kırıkçı, K. (2017). Effects of different storage time on hatching results and some egg quality characteristics of rock partridge (*A. graeca*) (management and production). *Poultry Science*, 96(6), 1628–1634. <https://doi.org/https://doi.org/10.3382/ps/pew443>

Haugh, R. R. (1937). The Haugh unit for measuring egg quality. *United States Egg and Poultry Magazine*, 43, 522–555.

He, X., Zhao, T., Shen, F., Liu, Q., Fang, Y., & Hu, Q. (2021). Online detection of naturally DON contaminated wheat grains from China using Vis-NIR spectroscopy and computer vision. *Biosystems Engineering*, 201, 1–10. <https://doi.org/10.1016/j.biosystemseng.2020.11.001>

Hourant, P., Baeten, V., Morales, M. T., Meurens, M., & Aparicio, R. (2000). Oil and Fat Classification by Selected Bands of Near-Infrared Spectroscopy. *Applied Spectroscopy*, 54(8), 1168–1174. <http://as.osa.org/abstract.cfm?URI=as-54-8-1168>

IBGE. (2020). *Pesquisa da Pecuária Municipal*. https://agenciadenoticias.ibge.gov.br/media/com_mediaibge/arquivos/b25372bebf621f8c789c4fda346d1a9.pdf

Jiang, S., Zhu, T., Jia, D., Yao, J., & Jiang, Y. (2018). Effect of egg freshness on their automatic orientation. *Journal of the Science of Food and Agriculture*, 98(7), 2642–2650. <https://doi.org/10.1002/jsfa.8757>

Kaufmann, K. C., Favero, F. de F., de Vasconcelos, M. A. M., Godoy, H. T., Sampaio, K. A., & Barbin, D. F. (2019). Portable NIR Spectrometer for Prediction of Palm Oil Acidity. *Journal of Food Science*, 84(3), 406–411. <https://doi.org/https://doi.org/10.1111/1750-3841.14467>

Kumbár, V., Trnka, J., Nedomová, Š., & Buchar, J. (2015). On the influence of storage duration on rheological properties of liquid egg products and response of eggs to impact loading - Japanese quail eggs. *Journal of Food Engineering*, 166, 86–94. <https://doi.org/10.1016/j.jfoodeng.2015.05.030>

Lin, H., Zhao, J., Sun, L., Chen, Q., & Zhou, F. (2011). Freshness measurement of eggs using near infrared (NIR) spectroscopy and multivariate data analysis. *Innovative Food Science & Emerging Technologies*, 12(2), 182–186. <https://doi.org/https://doi.org/10.1016/j.ifset.2011.01.008>

Liu, L., Yang, R., Luo, X., Dong, K., Huang, X., Song, H., Gao, H., Li, S., & Huang, Q. (2020). Omics analysis of holoproteins and modified proteins of quail egg. *Food Chemistry*, 326(April), 126983. <https://doi.org/10.1016/j.foodchem.2020.126983>

Marinho, A. L. (2011). *Qualidade interna e externa de ovos de codornas japonesas armazenados em diferentes temperaturas e períodos de estocagem*.

McVey, C., McGrath, T. F., Haughey, S. A., & Elliott, C. T. (2021). A rapid food chain approach for authenticity screening: The development, validation and transferability of a chemometric model using two handheld near infrared spectroscopy (NIRS) devices. *Talanta*, 222, 121533. <https://doi.org/https://doi.org/10.1016/j.talanta.2020.121533>

Mendez, J., Mendoza, L., Cruz-Tirado, J. P., Quevedo, R., & Siche, R. (2019). Trends in application of NIR and hyperspectral imaging for food authentication . In *Scientia Agropecuaria* (Vol. 10, pp. 143–161). scielo .

Narushin, V. G., Romanov, M. N., & Griffin, D. K. (2021). A novel Egg Quality Index as an alternative to Haugh unit score. *Journal of Food Engineering*, 289(March 2020), 110176. <https://doi.org/10.1016/j.jfoodeng.2020.110176>

Páscoa, R. N. M. J., Magalhães, L. M., & Lopes, J. A. (2013). FT-NIR spectroscopy as a tool for valorization of spent coffee grounds: Application to assessment of antioxidant properties. *Food Research International*, 51(2), 579–586. <https://doi.org/https://doi.org/10.1016/j.foodres.2013.01.035>

Pereira, E. V. dos S., Fernandes, D. D. de S., de Araújo, M. C. U., Diniz, P. H. G. D., & Maciel, M. I. S. (2020). Simultaneous determination of goat milk adulteration with cow milk and their fat and protein contents using NIR spectroscopy and PLS algorithms. *Lwt*, 127(April), 109427. <https://doi.org/10.1016/j.lwt.2020.109427>

Saey, W., Mouazen, A. M., & Ramon, H. (2005). Potential for Onsite and Online Analysis of Pig Manure using Visible and Near Infrared Reflectance Spectroscopy. *Biosystems Engineering*, 91(4), 393–402.

<https://doi.org/https://doi.org/10.1016/j.biosystemseng.2005.05.001>

Samiullah, S., & Roberts, J. R. (2014). The eggshell cuticle of the laying hen. *World's Poultry Science Journal*, 70(4), 693–708. [https://doi.org/DOI: 10.1017/S0043933914000786](https://doi.org/DOI:10.1017/S0043933914000786)

Sari, M., Tilki, M., & Saatci, M. (2016). Genetic parameters of egg quality traits in long-term pedigree recorded Japanese quail. *Poultry Science*, 95(8), 1743–1749. **<https://doi.org/10.3382/ps/pew118>**

Sato, T., Kawano, S., & Iwamoto, M. (1991). Near infrared spectral patterns of fatty acid analysis from fats and oils. *Journal of the American Oil Chemists Society*, 68(11), 827–833. **<https://doi.org/10.1007/BF02660596>**

Sharp, P. F., & Powell, C. K. (1930). Decrease in Interior Quality of Hens' Eggs during Storage as Indicated by the Yolk. *Industrial & Engineering Chemistry*, 22(8), 908–910. **<https://doi.org/10.1021/ie50248a031>**

Sun, C., Liu, J., Yang, N., & Xu, G. (2017). Egg quality and egg albumen property of domestic chicken , duck , goose , turkey , quail , and pigeon. *Poultry Science*, 98(10), 4516–4521. **<https://doi.org/10.3382/ps/pez259>**

Vapnik, V. (2013). *The nature of statistical learning theory*. Springer science & business media.

Walsh, K. B., Blasco, J., Zude-Sasse, M., & Sun, X. (2020). Visible-NIR ‘point’ spectroscopy in postharvest fruit and vegetable assessment: The science behind three decades of commercial use. *Postharvest Biology and Technology*, 168, 111246. **<https://doi.org/https://doi.org/10.1016/j.postharvbio.2020.111246>**

ang, S., Chen, W., Zhang, H. X., Ruan, D., & Lin, Y. C. (2014). Influence of particle size and calcium source on production performance, egg quality, and bone parameters in laying ducks. *Poultry Science*, 93(10), 2560–2566. **<https://doi.org/https://doi.org/10.3382/ps.2014-03962>**

Ziegler, J. U., Leitenberger, M., Longin, C. F. H., Würschum, T., Carle, R., & Schweiggert, R. M. (2016). Near-infrared reflectance spectroscopy for the rapid discrimination of kernels and flours of different wheat species. *Journal of Food Composition and Analysis*, 51, 30–36. **<https://doi.org/https://doi.org/10.1016/j.jfca.2016.06.005>**

CAPÍTULO 4:
CONCLUSÃO GERAL

CONCLUSÃO GERAL

Esta dissertação demonstra a efetividade da espectroscopia NIR, combinada com os algoritmos PLS e SVM, para avaliação não destrutiva da qualidade interna e do frescor de ovos de codorna através de sensores portáteis e manuais de baixo custo e de fácil utilização. Os modelos de regressão obtidos para o SVM forneceram uma melhor precisão para todos os elementos, quando comparado aos modelos PLS, já os modelos de classificação apresentaram desempenho bastante similar para ambos os algoritmos. Vale ressaltar que a predição dos parâmetros HU, YI e EQI para avaliação da qualidade e do frescor de ovos de codorna por espectroscopia NIR são relatados pela primeira vez na literatura, viabilizando a aplicação desta técnica na indústria de ovoprodutos, uma vez que esta pode ser empregada em toda a cadeia produtiva de forma online.

CAPÍTULO 5:

REFERÊNCIAS

REFERÊNCIAS

- ABU-KHALAF, N.; HMIDAT, M. Visible/Near Infrared (VIS/NIR) spectroscopy as an optical sensor for evaluating olive oil quality. **Computers and Electronics in Agriculture**, v. 173, p. 105445, 2020.
- BALCEROWSKA-CZERNIAK, G. et al. Study of surface oxidation of polycrystalline rhodium using multivariate analysis based on combined optical microscopy and X-ray photoelectron spectroscopy data. **Applied Surface Science**, v. 526, n. May, p. 146617, 2020.
- BAO, Z. et al. Effect of salting on the water migration, physicochemical and textural characteristics, and microstructure of quail eggs. **LWT**, v. 132, p. 109847, 2020.
- BARBOSA, T. C. G. PARÂMETROS DE QUALIDADE INTERNA E EXTERNA DE OVOS DE CODORNA. 2013.
- BARRAGÁN, W. et al. Authentication of barley-finished beef using visible and near infrared spectroscopy (Vis-NIRS) and different discrimination approaches. **Meat Science**, v. 172, p. 108342, 2021.
- BEDIN, F. C. B. et al. NIR associated to PLS and SVM for fast and non-destructive determination of C, N, P, and K contents in poultry litter. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, v. 245, p. 118834, 2021.
- BRAGAGNOLO, N.; RODRIGUEZ-AMAYA, D. B. Comparison of the cholesterol content of Brazilian chicken and quail eggs. **Journal of Food Composition and Analysis**, v. 16, n. 2, p. 147–153, 2003.
- BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Portaria nº 1, de 21 de fevereiro de 1990.
- DE SOUZA, A. M.; POPPI, R. J. Teaching experiment of chemometrics for exploratory analysis of edible vegetable oils by mid infrared spectroscopy and principal component analysis: A tutorial, part I. **Química Nova**, v. 35, n. 1, p. 223–229, 2012.
- DEVOS, O. et al. Support vector machines (SVM) in near infrared (NIR) spectroscopy: Focus on parameters optimization and model interpretation. **Chemometrics and Intelligent Laboratory Systems**, v. 96, n. 1, p. 27–33, 2009.
- EL-TARABANY, M. S. Effect of thermal stress on fertility and egg quality of Japanese quail. **Journal of Thermal Biology**, v. 61, p. 38–43, 2016.

ESQUERRE, C. A. et al. Use of an NIR MEMS spectrophotometer and visible/NIR hyperspectral imaging systems to predict quality parameters of treated ground peppercorns. **LWT**, v. 131, p. 109761, 2020.

FERREIRA, D. Structural imaging in dementia with Lewy bodies: the potential of multivariate data analysis. **Psychiatry Research - Neuroimaging**, v. 306, n. June, p. 111180, 2020.

GIUNCHI, A. et al. Non-destructive freshness assessment of shell eggs using FT-NIR spectroscopy. **Journal of Food Engineering**, v. 89, n. 2, p. 142–148, 2008.

GODOY, L. B. L. N. DE. Determinação de parâmetros de qualidade em madeira utilizando espectroscopia no infravermelho próximo e calibração multivariada. 2010.

HE, X. et al. Online detection of naturally DON contaminated wheat grains from China using Vis-NIR spectroscopy and computer vision. **Biosystems Engineering**, v. 201, p. 1–10, 2021.

IBGE. **Pesquisa da Pecuária Municipal**. [s.l: s.n.]. Disponível em: <https://agenciadenoticias.ibge.gov.br/media/com_mediaibge/arquivos/b25372bebf621f8c789c4fda346d1a9.pdf>.

KEMSLEY, E. K.; DEFERNEZ, M.; MARINI, F. Multivariate statistics: Considerations and confidences in food authenticity problems. **Food Control**, v. 105, p. 102–112, 2019.

KUMBÁR, V. et al. On the influence of storage duration on rheological properties of liquid egg products and response of eggs to impact loading – Japanese quail eggs. **Journal of Food Engineering**, v. 166, p. 86–94, 2015.

LIU, L. et al. Omics analysis of holoproteins and modified proteins of quail egg. **Food Chemistry**, v. 326, n. April, p. 126983, 2020.

LOFFREDI, E.; GRASSI, S.; ALAMPRESE, C. Spectroscopic approaches for non-destructive shell egg quality and freshness evaluation: Opportunities and challenges. **Food Control**, v. 129, p. 108255, 2021.

MALFATTI, L. H. et al. Evaluating and predicting egg quality indicators through principal component analysis and artificial neural networks. **LWT**, v. 148, p. 111720, 2021.

MARETTO, D. A. Aplicação de máquinas de vetores de suporte para desenvolvimento de modelos de classificação e calibração multivariada em espectroscopia no infravermelho. 2011.

MARINHO, A. L. Qualidade interna e externa de ovos de codornas japonesas armazenados em diferentes temperaturas e períodos de estocagem. 2011.

MARQUES, E. J. N. et al. Espectroscopia na região do infravermelho próximo (NIR): técnica analítica não destrutiva para determinação da qualidade de manga. In: **Manga: maturação, colheita e conservação**. [s.l: s.n.]. p. 63–87.

MENDOZA, D. et al. FT-NIR spectroscopy and RP-HPLC combined with multivariate analysis reveals differences in plant cell suspension cultures of *Thevetia peruviana* treated with salicylic acid and methyl jasmonate. **Biotechnology Reports**, v. 27, p. e00519, 2020.

MIGUEL, M. et al. Comparative study of egg white proteins from different species by chromatographic and electrophoretic methods. **European Food Research and Technology**, v. 221, n. 3, p. 542–546, 2005.

MORGANO, M. A. et al. Determinação de açúcar total em café cru por espectroscopia no infravermelho próximo e regressão por mínimos quadrados parciais. **Química Nova**, v. 30, n. 2, p. 346–350, 2007.

MÓRI, C. et al. Performance and egg quality of four quail genetic groups. **Revista Brasileira de Zootecnia**, v. 34, n. 3, p. 864–869, 2005.

NARUSHIN, V. G.; ROMANOV, M. N.; GRIFFIN, D. K. A novel Egg Quality Index as an alternative to Haugh unit score. **Journal of Food Engineering**, v. 289, n. March 2020, p. 110176, 2021.

NASR, M. A. F. et al. Does light intensity affect the behavior, welfare, performance, meat quality, amino acid profile, and egg quality of Japanese quails? **Poultry Science**, v. 98, n. 8, p. 3093–3102, 2019.

NIAZI, A.; GOODARZI, M.; YAZDANIPOUR, A. A comparative study between least-squares support vector machines and partial least squares in simultaneous spectrophotometric determination of cypermethrin, permethrin and tetramethrin. **Journal of the Brazilian Chemical Society**, v. 19, n. 3, p. 536–542, 2008.

OLADIPO, G. O. et al. Quail (*Coturnix japonica*) egg attenuated 2-butoxyethanol-induced enzymatic dysregulation, disseminated thrombosis and hemolytic impairment in female wistar rats. **Journal of Ethnopharmacology**, n. May, 2020.

OLIVEIRA, B. L. DE; OLIVEIRA, D. D. DE. Qualidade e tecnologia de ovos. **Lavras: Editora UFLA (Universidade Federal de Lavras)**, p. 223, 2013.

PIRES, P. G. S. et al. Effects of rice protein coating enriched with essential oils on internal quality and shelf life of eggs during room temperature storage. **Poultry Science**, v. 99, n. 1, p.

604–611, 2019.

PUERTAS, G.; VÁZQUEZ, M. UV-VIS-NIR spectroscopy and artificial neural networks for the cholesterol quantification in egg yolk. **Journal of Food Composition and Analysis**, v. 86, n. May 2019, 2020.

PULLANAGARI, R. R.; LI, M. Uncertainty assessment for firmness and total soluble solids of sweet cherries using hyperspectral imaging and multivariate statistics. **Journal of Food Engineering**, v. 289, n. January 2020, 2020.

RÉHAULT-GODBERT, S.; GUYOT, N.; NYS, Y. The Golden Egg: Nutritional Value, Bioactivities, and Emerging Benefits for Human Health. **Nutrients**, v. 11, n. 3, mar. 2019.

REN, Y. et al. Application of volatile and spectral profiling together with multimode data fusion strategy for the discrimination of preserved eggs. **Food Chemistry**, n. July, p. 128515, 2020.

RIGOLLE, A. et al. Development of an ultrasonic shear reflection technique to monitor the crystallization of cocoa butter. **Food Research International**, v. 75, p. 115–122, 2015.

ROSSI, G. B.; LOZANO, V. A. Simultaneous determination of quality parameters in yerba mate (*Ilex paraguariensis*) samples by application of near-infrared (NIR) spectroscopy and partial least squares (PLS). **LWT - Food Science and Technology**, v. 126, p. 109290, 2020.

SABIN, J. G.; FERRÃO, M. F.; FURTADO, J. C. Análise multivariada aplicada na identificação de fármacos antidepressivos. Parte II: Análise por componentes principais (PCA) e o método de classificação SIMCA. **Revista Brasileira de Ciências Farmacêuticas**, v. 40, n. 3, p. 387–396, 2004.

SANTOS, F. D.; SANTOS, L. P.; CUNHA, P. H.; BORGHI, F. T.; ROMÃO, W.; CASTRO, E. V. R.; OLIVEIRA, E. C.; FILGUEIRAS, P. R. Discrimination of oils and fuels using a portable NIR spectrometer. **Fuel**, v. 283, p. 118854, 2021.

SARI, M.; TILKI, M.; SAATCI, M. Genetic parameters of egg quality traits in long-term pedigree recorded Japanese quail. **Poultry Science**, v. 95, n. 8, p. 1743–1749, 2016.

SHARAF EDDIN, A.; IBRAHIM, S. A.; TAHERGORABI, R. Egg quality and safety with an overview of edible coating application for egg preservation. **Food Chemistry**, v. 296, p. 29–39, 2019.

SOHAIB ALI SHAH, S. et al. Towards fruit maturity estimation using NIR spectroscopy. **Infrared Physics and Technology**, v. 111, n. June, p. 103479, 2020.

SUN, C. et al. Egg quality and egg albumen property of domestic chicken, duck, goose, turkey, quail, and pigeon. **Poultry Science**, v. 98, n. 10, p. 4516–4521, 2019.

TERAMOTO, É. T. et al. Comparing different methods for estimating hourly solar ultraviolet radiation: Empirical models, artificial neural network and support vector machine. **Revista Brasileira de Meteorologia**, v. 35, n. 1, p. 35–43, 2020.

TEYE, E. et al. Cocoa bean and cocoa bean products quality evaluation by NIR spectroscopy and chemometrics: A review. **Infrared Physics and Technology**, v. 104, n. July 2019, 2020.

VALDERRAMA, L. et al. Proposal of a Didactic Experiment To Teach Principal Component Analysis. **Química Nova**, v. 39, n. 2, p. 245–249, 2015.